



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|--|--|--|--|---------|--------|----------------------|------|---------|---|--------|----------------------|------|---------|---|----------------------|-------------------------------|----|---------|----|--------------------|--------------|--|---------|----|--------------------|--------------|--|---------|----|--------------------------|--------------|----|---------|----|----------|-------|---------|--|----|--|-------|---------|--|----|-------------------------------------|-------|----|---------|----|-------|-------|---------|--|----|-------------------|---------------|---------|--|-----|------------------------------------|---|----|---------|----|-------|-------|---------|--|----|--------------------|----------|---------|--|-----|------------------------------------|-------|----|---------|----|-------|-------|---------|--|-----|-----------------|-------|---------|--|-----|--------------------------------------|-------|----|---------|
| (51) International Patent Classification 7 : C12N 15/12, 5/10, 1/21, C07K 14/705, 14/72, 16/28, A61K 38/17, C12Q 1/68 | | A2 | (11) International Publication Number: WO 00/20590 (43) International Publication Date: 13 April 2000 (13.04.00) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| (21) International Application Number: PCT/US99/23317 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| (22) International Filing Date: 6 October 1999 (06.10.99) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| (30) Priority Data: 09/167,219 6 October 1998 (06.10.98) US 60/172,211 6 October 1998 (06.10.98) US 60/133,585 11 May 1999 (11.05.99) US | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| (63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications US 09/167,219 (CIP) Filed on 6 October 1998 (06.10.98) US 60/172,211 (CIP) Filed on 6 October 1998 (06.10.98) US 60/133,585 (CIP) Filed on 11 May 1999 (11.05.99) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| (71) Applicant (for all designated States except US): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US). | | (74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US). | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| (72) Inventors; and (75) Inventors/Applicants (for US only): TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). YUE, | | (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <p>Published <i>Without international search report and to be republished upon receipt of that report.</i> </p> <p>(54) Title: G-PROTEIN COUPLED RECEPTOR PROTEINS</p> <table border="1"> <tr> <td>1</td> <td>M-----</td> <td>TPQSLLQTTLFLLSL-----</td> <td>LFLV</td> <td>2763296</td> </tr> <tr> <td>1</td> <td>M-----</td> <td>TPQSLLQTTLFLLSL-----</td> <td>LFLV</td> <td>1595762</td> </tr> <tr> <td>1</td> <td>MVFSVRQCGHVGRTBEVLLT</td> <td>F K I F L V I I C L H V V L V</td> <td>GI</td> <td>2117161</td> </tr> <tr> <td>21</td> <td>QGAHGRGHREDFR-----</td> <td>FCSQRNQTHRSS</td> <td></td> <td>2763296</td> </tr> <tr> <td>21</td> <td>QGAHGRGHREDFR-----</td> <td>FCSQRNQTHRSS</td> <td></td> <td>1595762</td> </tr> <tr> <td>36</td> <td>TSLEEDTDNSLSSPPPAKLSSVVS</td> <td>FAPSSNEVETTS</td> <td>GI</td> <td>2117161</td> </tr> <tr> <td>46</td> <td>LHY-----</td> <td>-----</td> <td>2763296</td> <td></td> </tr> <tr> <td>46</td> <td>LHYKPTEDLRISIENSEEALTVHAPPAAHPASR-----</td> <td>-----</td> <td>1595762</td> <td></td> </tr> <tr> <td>71</td> <td>LN-DVTLSSLPSNETEKTKTITIVKTFNAGVKPQR</td> <td>-----</td> <td>GI</td> <td>2117161</td> </tr> <tr> <td>49</td> <td>-----</td> <td>-----</td> <td>2763296</td> <td></td> </tr> <tr> <td>80</td> <td>[SPPDPGRLYHP-----</td> <td>[CLYWNRH-----</td> <td>1595762</td> <td></td> </tr> <tr> <td>105</td> <td>NICNLSSICNDSA[FRGBIMFQYDKESTVQPQNH</td> <td>I</td> <td>GI</td> <td>2117161</td> </tr> <tr> <td>49</td> <td>-----</td> <td>-----</td> <td>2763296</td> <td></td> </tr> <tr> <td>98</td> <td>[AGRLHLLYGLKD-----</td> <td>[FL-----</td> <td>1595762</td> <td></td> </tr> <tr> <td>140</td> <td>TNGTLTGVLSSLELKRSBELNKTQLSETYFIMCA</td> <td>-----</td> <td>GI</td> <td>2117161</td> </tr> <tr> <td>49</td> <td>-----</td> <td>-----</td> <td>2763296</td> <td></td> </tr> <tr> <td>113</td> <td>[SDKASSLLC-----</td> <td>-----</td> <td>1595762</td> <td></td> </tr> <tr> <td>175</td> <td>TABAQISTLNCTFTIKLNNTMNACAAIAALERVKIR</td> <td>-----</td> <td>GI</td> <td>2117161</td> </tr> </table> <p>(57) Abstract</p> <p>The invention provides human G-protein coupled receptor proteins (GCRP) and polynucleotides which identify and encode GCRP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of GCRP.</p> | | | | 1 | M----- | TPQSLLQTTLFLLSL----- | LFLV | 2763296 | 1 | M----- | TPQSLLQTTLFLLSL----- | LFLV | 1595762 | 1 | MVFSVRQCGHVGRTBEVLLT | F K I F L V I I C L H V V L V | GI | 2117161 | 21 | QGAHGRGHREDFR----- | FCSQRNQTHRSS | | 2763296 | 21 | QGAHGRGHREDFR----- | FCSQRNQTHRSS | | 1595762 | 36 | TSLEEDTDNSLSSPPPAKLSSVVS | FAPSSNEVETTS | GI | 2117161 | 46 | LHY----- | ----- | 2763296 | | 46 | LHYKPTEDLRISIENSEEALTVHAPPAAHPASR----- | ----- | 1595762 | | 71 | LN-DVTLSSLPSNETEKTKTITIVKTFNAGVKPQR | ----- | GI | 2117161 | 49 | ----- | ----- | 2763296 | | 80 | [SPPDPGRLYHP----- | [CLYWNRH----- | 1595762 | | 105 | NICNLSSICNDSA[FRGBIMFQYDKESTVQPQNH | I | GI | 2117161 | 49 | ----- | ----- | 2763296 | | 98 | [AGRLHLLYGLKD----- | [FL----- | 1595762 | | 140 | TNGTLTGVLSSLELKRSBELNKTQLSETYFIMCA | ----- | GI | 2117161 | 49 | ----- | ----- | 2763296 | | 113 | [SDKASSLLC----- | ----- | 1595762 | | 175 | TABAQISTLNCTFTIKLNNTMNACAAIAALERVKIR | ----- | GI | 2117161 |
| 1 | M----- | TPQSLLQTTLFLLSL----- | LFLV | 2763296 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | M----- | TPQSLLQTTLFLLSL----- | LFLV | 1595762 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | MVFSVRQCGHVGRTBEVLLT | F K I F L V I I C L H V V L V | GI | 2117161 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 21 | QGAHGRGHREDFR----- | FCSQRNQTHRSS | | 2763296 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 21 | QGAHGRGHREDFR----- | FCSQRNQTHRSS | | 1595762 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 36 | TSLEEDTDNSLSSPPPAKLSSVVS | FAPSSNEVETTS | GI | 2117161 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 46 | LHY----- | ----- | 2763296 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 46 | LHYKPTEDLRISIENSEEALTVHAPPAAHPASR----- | ----- | 1595762 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 71 | LN-DVTLSSLPSNETEKTKTITIVKTFNAGVKPQR | ----- | GI | 2117161 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 49 | ----- | ----- | 2763296 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 80 | [SPPDPGRLYHP----- | [CLYWNRH----- | 1595762 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 105 | NICNLSSICNDSA[FRGBIMFQYDKESTVQPQNH | I | GI | 2117161 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 49 | ----- | ----- | 2763296 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 98 | [AGRLHLLYGLKD----- | [FL----- | 1595762 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 140 | TNGTLTGVLSSLELKRSBELNKTQLSETYFIMCA | ----- | GI | 2117161 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 49 | ----- | ----- | 2763296 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 113 | [SDKASSLLC----- | ----- | 1595762 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 175 | TABAQISTLNCTFTIKLNNTMNACAAIAALERVKIR | ----- | GI | 2117161 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | | | |
|----|--------------------------|----|---------------------------------------|----|---|----|--------------------------|
| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| AT | Austria | FR | France | LU | Luxembourg | SN | Senegal |
| AU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav Republic of Macedonia | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | ML | Mali | TR | Turkey |
| BG | Bulgaria | HU | Hungary | MN | Mongolia | TT | Trinidad and Tobago |
| BJ | Benin | IE | Ireland | MR | Mauritania | UA | Ukraine |
| BR | Brazil | IL | Israel | MW | Malawi | UG | Uganda |
| BY | Belarus | IS | Iceland | MX | Mexico | US | United States of America |
| CA | Canada | IT | Italy | NE | Niger | UZ | Uzbekistan |
| CF | Central African Republic | JP | Japan | NL | Netherlands | VN | Viet Nam |
| CG | Congo | KE | Kenya | NO | Norway | YU | Yugoslavia |
| CH | Switzerland | KG | Kyrgyzstan | NZ | New Zealand | ZW | Zimbabwe |
| CI | Côte d'Ivoire | KP | Democratic People's Republic of Korea | PL | Poland | | |
| CM | Cameroon | KR | Republic of Korea | PT | Portugal | | |
| CN | China | KZ | Kazakhstan | RO | Romania | | |
| CU | Cuba | LC | Saint Lucia | RU | Russian Federation | | |
| CZ | Czech Republic | LI | Liechtenstein | SD | Sudan | | |
| DE | Germany | LK | Sri Lanka | SE | Sweden | | |
| DK | Denmark | LR | Liberia | SG | Singapore | | |
| EE | Estonia | | | | | | |

G-PROTEIN COUPLED RECEPTOR PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of G-protein coupled receptor
5 proteins and to the use of these sequences in the diagnosis, treatment, and prevention of disorders of the nervous system, autoimmune/inflammatory disorders, and cell proliferative disorders including cancer.

BACKGROUND OF THE INVENTION

Regulation of cell proliferation, differentiation, and migration is important for the formation and
10 function of tissues. Regulatory proteins such as growth factors control these cellular processes and act as mediators in cell-cell signaling pathways. Growth factors are secreted proteins that bind to specific cell-surface receptors on target cells. The bound receptors trigger intracellular signal transduction pathways which activate various downstream effectors that regulate gene expression, cell division, cell differentiation, cell motility, and other cellular processes. Some of the receptors involved in signal
15 transduction by growth factors belong to the large superfamily of G-protein coupled receptors (GPCRs).

GPCRs are integral membrane proteins characterized by the presence of seven hydrophobic transmembrane domains which together form a bundle of antiparallel alpha (α) helices. These proteins range in size from under 400 to over 1000 amino acids (Strosberg, A.D. (1991) Eur. J. Biochem. 196:1-10; Coughlin, S.R. (1994) Curr. Opin. Cell Biol. 6:191-197). The amino-terminus of a GPCR is
20 extracellular, is of variable length, and is often glycosylated. The carboxy-terminus is cytoplasmic and generally phosphorylated. Extracellular loops of GPCRs alternate with intracellular loops and link the transmembrane domains. Cysteine disulfide bridges linking the second and third extracellular loops may interact with agonists and antagonists. The most conserved domains of GPCRs are the transmembrane domains and the first two cytoplasmic loops. The transmembrane domains account for structural and
25 functional features of the receptor. In most cases, the bundle of α helices forms a ligand-binding pocket. The extracellular N-terminal segment, or one or more of the three extracellular loops, may also participate in ligand binding. Ligand binding activates the receptor by inducing a conformational change in intracellular portions of the receptor. In turn, the large, third intracellular loop of the activated receptor interacts with an intracellular heterotrimeric guanine nucleotide binding (G) protein complex which
30 mediates further intracellular signaling activities, including the activation of second messengers such as cyclic AMP (cAMP), phospholipase C, inositol triphosphate, or ion channel proteins. (See, e.g., Watson, S. and S. Arkinstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego CA, pp. 2-6; Bolander, F.F. (1994) Molecular Endocrinology, Academic Press, San Diego CA, pp. 162-176; Baldwin, J.M. (1994) Curr. Opin. Cell Biol. 6:180-190.)

35 GPCRs include receptors for sensory signal mediators (e.g., light and olfactory stimulatory

molecules); adenosine, bombesin, bradykinin, endothelin, γ -aminobutyric acid (GABA), hepatocyte growth factor, melanocortins, neuropeptide Y, opioid peptides, opsins, somatostatin, tachykinins, vasoactive intestinal polypeptide family, and vasopressin; biogenic amines (e.g., dopamine, epinephrine and norepinephrine, histamine, glutamate (metabotropic effect), acetylcholine (muscarinic effect), and 5 serotonin); chemokines; lipid mediators of inflammation (e.g., prostaglandins and prostanoids, platelet activating factor, and leukotrienes); and peptide hormones (e.g., calcitonin, C₅a anaphylatoxin, follicle-stimulating hormone (FSH), gonadotropin-releasing hormone (GnRH), neurokinin, and thyrotropin-releasing hormone (TRH), and oxytocin). GPCRs which act as receptors for stimuli that have yet to be identified are known as orphan receptors.

10 GPCRs are implicated in inflammation and the immune response, and include the EGF module-containing, mucin-like hormone receptor (Emr1) and CD97 β receptor proteins. These receptors contain between three and seven potential calcium-binding EGF-like motifs (Baud, V. et al. (1995) Genomics 26:334-344; Gray, J.X. et al. (1996) J. Immunol. 157:5438-5447). These GPCRs are members of the recently characterized EGF-TM7 receptors family. In addition, post-translational modification of aspartic 15 acid or asparagine to form erythro- β -hydroxyaspartic acid or erythro- β -hydroxyasparagine has been identified in a number of proteins with domains homologous to EGF. The consensus pattern is located in the N-terminus of the EGF-like domain. Examples of such proteins are blood coagulation factors VII, IX, and X; proteins C, S, and Z; the LDL receptor; and thrombomodulin.

One large subfamily of GPCRs are the olfactory receptors. These receptors share the seven 20 hydrophobic transmembrane domains of other GPCRs and function by registering G protein-mediated transduction of odorant signals. Numerous distinct olfactory receptors are required to distinguish different odors. Each olfactory sensory neuron expresses only one type of olfactory receptor, and distinct spatial zones of neurons expressing distinct receptors are found in nasal pasages. One olfactory receptor, the RA1c receptor which was isolated from a rat brain library, has been shown to be limited in expression 25 to very distinct regions of the brain and a defined zone of the olfactory epithelium (Raming, K. et al. (1998) Receptors Channels 6:141-151). In another example, three rat genes encoding olfactory-like receptors having typical GPCR characteristics showed expression patterns exclusively in taste, olfactory, and male reproductive tissue (Thomas, M.B. et al. (1996) Gene 178:1-5).

Another group of GPCRs are the mas oncogene-related proteins. Like the mas oncogenes 30 themselves, some of these mas-like receptors are implicated in intracellular angiotensin II actions. Angiotensin II, an octapeptide hormone, mediates vasoconstriction and aldosterone secretion through angiotensin II receptor molecules found on smooth vascular muscle and the adrenal glands, respectively. A cloned human mas-related gene (mrg) mRNA, when injected into Xenopus oocytes, produces an 35 increase in the response to angiotensin peptides. Mrg has been shown to directly affect signaling pathways associated with the angiotensin II receptor, and, accordingly, affects the processes of

vasoconstriction and aldosterone secretion (Monnot, C. et al. (1991) Mol. Endocrinol. 5:1477-1487).

5 GPCR mutations, which may cause loss of function or constitutive activation, have been associated with numerous human diseases (Coughlin, *supra*). For instance, retinitis pigmentosa may arise from mutations in the rhodopsin gene. Rhodopsin is the retinal photoreceptor which is located within the discs of the eye rod cell. Parma, J. et al. (1993, Nature 365:649-651) reported that somatic activating mutations in the thyrotropin receptor cause hyperfunctioning thyroid adenomas and suggested that certain GPCRs susceptible to constitutive activation may behave as protooncogenes.

10 The discovery of new G-protein coupled receptor proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of disorders of the nervous system, autoimmune/inflammatory disorders, and cell proliferative disorders including cancer.

SUMMARY OF THE INVENTION

15 The invention features substantially purified polypeptides, G-protein coupled receptor proteins, referred to collectively as "GCRP" and individually as "GCRP-1," "GCRP-2," "GCRP-3," "GCRP-4," "GCRP-5," "GCRP-6," "GCRP-7," and "GCRP-8." In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-8, and fragments thereof.

20 The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-8, and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-8, and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising 25 an amino acid sequence selected from the group consisting of SEQ ID NO:1-8, and fragments thereof.

30 Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-8, and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-8, and fragments thereof.

35 The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of: (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates

with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:9-16, and fragments thereof. The invention 5 further provides an isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:9-16, and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:9-16, and fragments thereof.

10 The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-8, and fragments thereof. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the 15 steps of: (a) culturing the host cell containing an expression vector containing at least a fragment of a polynucleotide under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-8, and 20 fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-8, and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with 25 decreased expression or activity of GCRP, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-8, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with 30 increased expression or activity of GCRP, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-8, and fragments thereof.

BRIEF DESCRIPTION OF THE FIGURES AND TABLES

35 Figures 1A, 1B, 1C, 1D, and 1E show the amino acid sequence alignment between GCRP-2

(2763296; SEQ ID NO:2), GCRP-1 (1595762; SEQ ID NO:1), and human epididymis-specific seven transmembrane receptor HE6 (GI 2117161; SEQ ID NO:17), produced using the multisequence alignment program of LASERGENE software (DNASTAR, Madison WI).

Figures 2A and 2B show the amino acid sequence alignment between GCRP-3 (3367641; SEQ 5 ID NO:3) and human CD97 (GI 3342740; SEQ ID NO:18), produced using the multisequence alignment program of LASERGENE software.

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding GCRP.

10 Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods and algorithms used for identification of GCRP.

Table 3 shows useful fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

15 Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding GCRP were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze GCRP, along with applicable descriptions, references, and threshold parameters.

20 DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be 25 limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

30 Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, 35 protocols, reagents and vectors which are reported in the publications and which might be used in

connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

“GCRP” refers to the amino acid sequences of substantially purified GCRP obtained from any 5 species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and preferably the human species, from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term “agonist” refers to a molecule which, when bound to GCRP, increases or prolongs the duration of the effect of GCRP. Agonists may include proteins, nucleic acids, carbohydrates, or any other 10 molecules which bind to and modulate the effect of GCRP.

An “allelic variant” is an alternative form of the gene encoding GCRP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in 15 polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

“Altered” nucleic acid sequences encoding GCRP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polynucleotide the same as GCRP or a 20 polypeptide with at least one functional characteristic of GCRP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding GCRP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding GCRP. The encoded protein may also be “altered,” and may contain deletions, insertions, or substitutions of amino 25 acid residues which produce a silent change and result in a functionally equivalent GCRP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of GCRP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, positively charged amino acids may include lysine and arginine, and 30 amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; and phenylalanine and tyrosine.

The terms “amino acid” and “amino acid sequence” refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. In 35 this context, “fragments,” “immunogenic fragments,” or “antigenic fragments” refer to fragments of

GCRP which are preferably at least 5 to about 15 amino acids in length, most preferably at least 14 amino acids, and which retain some biological activity or immunological activity of GCRP. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native

5 amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence.

Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which, when bound to GCRP, decreases the amount or 10 the duration of the effect of the biological or immunological activity of GCRP. Antagonists may include proteins, nucleic acids, carbohydrates, antibodies, or any other molecules which decrease the effect of GCRP.

The term "antibody" refers to intact molecules as well as to fragments thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding the epitopic determinant. Antibodies that bind 15 GCRP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet 20 hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that fragment of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind 25 specifically to antigenic determinants (given regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the 30 complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" can refer to the antisense strand, and the designation "positive" can refer to the sense strand.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability 35 of the natural, recombinant, or synthetic GCRP, or of any oligopeptide thereof, to induce a specific

immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" and "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such 5 that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid (PNA) molecules.

10 A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding GCRP or fragments of GCRP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated 15 with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' 20 direction, and resequenced, or which has been assembled from the overlapping sequences of more than one Incyte clone using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

The term "correlates with expression of a polynucleotide" indicates that the detection of the 25 presence of nucleic acids, the same or related to a nucleic acid sequence encoding GCRP, by northern analysis is indicative of the presence of nucleic acids encoding GCRP in a sample, and thereby correlates with expression of the transcript from the polynucleotide encoding GCRP.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

30 The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at 35 least one biological or immunological function of the polypeptide from which it was derived.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely 5 complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as 10 reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

15 The phrases "percent identity" and "% identity" refer to the percentage of sequence similarity found in a comparison of two or more amino acid or nucleic acid sequences. Percent identity can be determined electronically, e.g., by using the MEGALIGN program (DNASTAR, Madison WI) which creates alignments between two or more sequences according to methods selected by the user, e.g., the clustal method. (See, e.g., Higgins, D.G. and P.M. Sharp (1988) Gene 73:237-244.) Parameters for each 20 method may be the default parameters provided by MEGALIGN or may be specified by the user. The clustal algorithm groups sequences into clusters by examining the distances between all pairs. The clusters are aligned pairwise and then in groups. The percentage similarity between two amino acid sequences, e.g., sequence A and sequence B, is calculated by dividing the length of sequence A, minus the number of gap residues in sequence A, minus the number of gap residues in sequence B, into the sum 25 of the residue matches between sequence A and sequence B, times one hundred. Gaps of low or of no similarity between the two amino acid sequences are not included in determining percentage similarity. Percent identity between nucleic acid sequences can also be counted or calculated by other methods known in the art, e.g., the Jotun Hein method. (See, e.g., Hein, J. (1990) Methods Enzymol. 183:626-645.) Identity between sequences can also be determined by other methods known in the art, e.g., by 30 varying hybridization conditions.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence 35 in the non-antigen binding regions has been altered so that the antibody more closely resembles a human

antibody, and still retains its original binding ability.

“Hybridization” refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

The term “hybridization complex” refers to a complex formed between two nucleic acid

5 sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C₆t or R₆t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

10 The words “insertion” and “addition” refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, to the sequence found in the naturally occurring molecule.

15 “Immune response” can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

20 The term “microarray” refers to an arrangement of distinct polynucleotides on a substrate.

The terms “element” and “array element” in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

25 The term “modulate” refers to a change in the activity of GCRP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of GCRP.

30 The phrases “nucleic acid” or “nucleic acid sequence,” as used herein, refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material. In this context, “fragments” refers to those nucleic acid sequences which comprise a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:9-16, for example, as distinct from any other sequence in the same genome. For example, a fragment of SEQ ID NO:9-16 is useful in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:9-16 from related polynucleotide sequences. A fragment of SEQ ID NO:9-16 is at least about 15-20 nucleotides in length. The precise length of the fragment of SEQ ID NO:9-16 and the region of SEQ ID NO:9-16 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment. In some cases, a fragment, when translated, would

35 produce polypeptides retaining some functional characteristic, e.g., antigenicity, or structural domain

characteristic, e.g., ATP-binding site, of the full-length polypeptide.

The terms "operably associated" and "operably linked" refer to functionally related nucleic acid sequences. A promoter is operably associated or operably linked with a coding sequence if the promoter controls the translation of the encoded polypeptide. While operably associated or operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements, e.g., repressor genes, are not contiguously linked to the sequence encoding the polypeptide but still bind to operator sequences that control expression of the polypeptide.

The term "oligonucleotide" refers to a nucleic acid sequence of at least about 6 nucleotides to 60 nucleotides, preferably about 15 to 30 nucleotides, and most preferably about 20 to 25 nucleotides, which 10 can be used in PCR amplification or in a hybridization assay or microarray. "Oligonucleotide" is substantially equivalent to the terms "amplimer," "primer," "oligomer," and "probe," as these terms are commonly defined in the art.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid 15 residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding GCRP, or fragments thereof, or GCRP itself, may comprise a bodily fluid; an extract from a 20 cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, or an antagonist. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the 25 binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "stringent conditions" refers to conditions which permit hybridization between polynucleotides and the claimed polynucleotides. Stringent conditions can be defined by salt 30 concentration, the concentration of organic solvent, e.g., formamide, temperature, and other conditions well known in the art. In particular, stringency can be increased by reducing the concentration of salt, increasing the concentration of formamide, or raising the hybridization temperature.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably 35 about 75% free, and most preferably about 90% free from other components with which they are naturally

associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, 5 slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well 10 known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or 15 as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of GCRP polypeptides refers to an amino acid sequence that is altered by one or more amino acid residues. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties (e.g., replacement of leucine with isoleucine). More rarely, a 20 variant may have "nonconservative" changes (e.g., replacement of glycine with tryptophan). Analogous minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, LASERGENE software (DNASTAR).

25 The term "variant," when used in the context of a polynucleotide sequence, may encompass a polynucleotide sequence related to GCRP. This definition may also include, for example, "allelic" (as defined above), "splice," "species," or "polymorphic" variants. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess 30 additional functional domains or an absence of domains. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one base. The 35 presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity

for a disease state.

THE INVENTION

The invention is based on the discovery of new human G-protein coupled receptor proteins (GCRP), the polynucleotides encoding GCRP, and the use of these compositions for the diagnosis, 5 treatment, or prevention of disorders of the nervous system, autoimmune/inflammatory disorders, and cell proliferative disorders including cancer.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding GCRP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic 10 acids encoding each GCRP were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The clones in column 5 were used to assemble the consensus nucleotide sequence of each GCRP and are useful as fragments in hybridization technologies.

15 The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs, as well as GenBank homologs identified by BLAST analysis; column 6 shows the identity of each polypeptide; and 20 column 7 shows analytical methods used to characterize each polypeptide through sequence homology and protein motifs.

As shown in Figures 1A, 1B, 1C, 1D, and 1E, GCRP-1 has chemical and structural similarity with 25 human epididymis-specific seven transmembrane receptor HE6 (GI 2117161; SEQ ID NO:17). In particular, GCRP-1 and human HE6 share 19% identity, four potential N-glycosylation sites, and one potential protein kinase C phosphorylation site. GCRP-1 and GCRP-2 have chemical and structural 30 homology and are considered to be splice variants. Therefore, GCRP-2 also has chemical and structural similarity with human epididymis-specific seven transmembrane receptor HE6 (GI 2117161; SEQ ID NO:17). As shown in Figures 2A and 2B, GCRP-3 has chemical and structural similarity with human CD97 (GI 3342740; SEQ ID NO:18). In particular, GCRP-3 and human CD97 share 44% identity, one potential N-glycosylation site, one potential protein kinase C phosphorylation site, and one potential aspartic acid/asparagine hydroxylation site.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding GCRP. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are 35 useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:9-16, and to

distinguish between SEQ ID NO:9-16 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express GCRP as a fraction of total tissues expressing GCRP. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing GCRP as a fraction of total tissues expressing GCRP. Column 5 lists the vectors used to subclone each cDNA library. Northern analysis shows the expression of SEQ ID NO:9 in various libraries, at least 55% of which are associated with cancer and cell proliferation, and at least 30% of which are associated with the immune response. Northern analysis shows the expression of SEQ ID NO:11 in various libraries, at least 50% of which are associated with cancer and at least 50% of which are associated with the immune response. Of particular note is the expression of SEQ ID NO:9 in brain, thyroid, epididymis, prostate, breast, uterus, kidney, lung, gut, and bladder; the expression of SEQ ID NO:10 in brain, kidney, epididymis, prostate, breast, and bladder; the expression of SEQ ID NO:11 in tumor, immunological, and inflammatory tissues; the expression of SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, and SEQ ID NO:16 in cancerous and proliferating tissues; the expression of SEQ ID NO:12 in brain tumor; the expression of SEQ ID NO:13 in kidney tumor and prostate tumor; and the expression of SEQ ID NO:14 in tongue tumor and prostate.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding GCRP were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

The invention also encompasses GCRP variants. A preferred GCRP variant is one which has at least about 80%, more preferably at least about 90%, and most preferably at least about 95% amino acid sequence identity to the GCRP amino acid sequence, and which contains at least one functional or structural characteristic of GCRP.

The invention also encompasses polynucleotides which encode GCRP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:9-16, which encodes GCRP.

The invention also encompasses a variant of a polynucleotide sequence encoding GCRP. In particular, such a variant polynucleotide sequence will have at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding GCRP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:9-16 which has at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:9-16. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of GCRP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding GCRP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made 5 by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring GCRP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode GCRP and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring GCRP under appropriately selected 10 conditions of stringency, it may be advantageous to produce nucleotide sequences encoding GCRP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding GCRP 15 and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode GCRP and GCRP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic 20 sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding GCRP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:9-16, and 25 fragments, thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the 30 absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and most preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion 35 or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are

accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30°C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37°C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 µg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, 5 hybridization will occur at 42°C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50 % formamide, and 200 µg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

The washing steps which follow hybridization can also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be 10 increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include temperature of at least about 25°C, more preferably of at least about 42°C, and most preferably of at least about 68°C. In a preferred embodiment, 15 wash steps will occur at 25°C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a most preferred embodiment, wash steps will occur at 68°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art.

20 Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system 25 (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA), and ABI CATALYST 800 thermal cycler (Perkin-Elmer). Sequencing is then carried out using the ABI 373 or 377 DNA sequencing systems (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The 30 resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding GCRP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such 35 as promoters and regulatory elements. For example, one method which may be employed, restriction-site

PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode GCRP may be cloned in recombinant DNA molecules that direct expression of GCRP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent

amino acid sequence may be produced and used to express GCRP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter GCRP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding GCRP may be synthesized, in whole or in part, 10 using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) *Nucleic Acids Symp. Ser.* 7:215-223; Horn, T. et al. (1980) *Nucleic Acids Symp. Ser.* 7:225-232.) Alternatively, GCRP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Perkin-15 Elmer). Additionally, the amino acid sequence of GCRP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) *Methods Enzymol.* 182:392-421.) The 20 composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) *Proteins, Structures and Molecular Properties*, WH Freeman, New York NY.)

In order to express a biologically active GCRP, the nucleotide sequences encoding GCRP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a 25 suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding GCRP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding GCRP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding 30 GCRP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The 35 efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host

cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding GCRP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo 5 genetic recombination. (See, e.g., Sambrook, J. et al. (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding GCRP. These include, but are not limited to, microorganisms such as bacteria transformed with 10 recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

15 In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding GCRP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding GCRP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding GCRP into the vector's multiple cloning site 20 disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509.) When large quantities of GCRP are needed, e.g. for the production of antibodies, 25 vectors which direct high level expression of GCRP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of GCRP. A number of vectors containing 30 constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, *supra*; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184.)

Plant systems may also be used for expression of GCRP. Transcription of sequences encoding 35 GCRP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311).

Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g.,

5 The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding GCRP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader 10 sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses GCRP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

15 Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of 20 GCRP in cell lines is preferred. For example, sequences encoding GCRP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a 25 selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase 30 genes, for use in *tk* or *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. 35 USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable

genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only 5 to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding GCRP is inserted within a marker gene sequence, transformed cells containing 10 sequences encoding GCRP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding GCRP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding GCRP and that express 15 GCRP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of GCRP using either 20 specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on GCRP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) 25 Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or 30 PCR probes for detecting sequences related to polynucleotides encoding GCRP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding GCRP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 35 and labeled nucleotides. These procedures may be conducted using a variety of commercially available

kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

5 Host cells transformed with nucleotide sequences encoding GCRP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode GCRP may be designed to contain signal sequences which direct secretion
10 of GCRP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may
15 also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

20 In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding GCRP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric GCRP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of GCRP activity. Heterologous protein and peptide moieties
25 may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-*
30 *myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the GCRP encoding sequence and the heterologous protein sequence, so that GCRP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are
35 discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to

facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled GCRP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract systems (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or 5 SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, preferably ³⁵S-methionine.

Fragments of GCRP may be produced not only by recombinant production, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for 10 example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of GCRP may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of GCRP and G-protein coupled receptor proteins. In addition, the expression of GCRP is closely 15 associated with cell proliferation, cancer, and autoimmune disease/inflammation. Therefore, GCRP appears to play a role in disorders of the nervous system, autoimmune/inflammatory disorders, and cell proliferative disorders including cancer. In the treatment of disorders associated with increased GCRP expression or activity, it is desirable to decrease the expression or activity of GCRP. In the treatment of disorders associated with decreased GCRP expression or activity, it is desirable to increase the expression 20 or activity of GCRP.

Therefore, in one embodiment, GCRP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GCRP.

Examples of such disorders include, but are not limited to, a nervous disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's 25 disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and 30 Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorder of the central nervous system, cerebral palsy, a neuroskeletal disorder, an autonomic nervous system disorder, a cranial nerve disorder, a spinal cord disease, muscular dystrophy and other neuromuscular disorders, a peripheral 35 nervous system disorder, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic

myopathy; myasthenia gravis, periodic paralysis; a mental disorder including mood, anxiety, and schizophrenic disorders; seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, 5 cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, 10 prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; and an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact 15 dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's 20 syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma.

In another embodiment, a vector capable of expressing GCRP or a fragment or derivative thereof 25 may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GCRP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified GCRP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GCRP including, but not limited to, those 30 provided above.

In still another embodiment, an agonist which modulates the activity of GCRP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GCRP including, but not limited to, those listed above.

In a further embodiment, an antagonist of GCRP may be administered to a subject to treat or 35 prevent a disorder associated with increased expression or activity of GCRP. Examples of such disorders

include, but are not limited to, disorders of the nervous system, autoimmune/inflammatory disorders, and cell proliferative disorders. In one aspect, an antibody which specifically binds GCRP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express GCRP.

5 In an additional embodiment, a vector expressing the complement of the polynucleotide encoding GCRP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of GCRP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate

10 therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

15 An antagonist of GCRP may be produced using methods which are generally known in the art. In particular, purified GCRP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind GCRP. Antibodies to GCRP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab 20 expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with GCRP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase 25 immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to GCRP 30 have an amino acid sequence consisting of at least about 5 amino acids, and, more preferably, of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of GCRP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

35 Monoclonal antibodies to GCRP may be prepared using any technique which provides for the

production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) *Nature* 256:495-497; Kozbor, D. et al. (1985) *J. Immunol. Methods* 81:31-42; Cote, R.J. et al. (1983) *Proc. Natl. Acad. Sci. USA* 80:2026-2030; and Cole, S.P. et al. (1984)

5 *Mol. Cell Biol.* 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) *Proc. Natl. Acad. Sci. USA* 81:6851-6855; Neuberger, M.S. et al. (1984) *Nature* 312:604-608; and Takeda, S. et al. (1985)

10 *Nature* 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce GCRP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) *Proc. Natl. Acad. Sci. USA* 88:10134-10137.)

15 Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:3833-3837; Winter, G. et al. (1991) *Nature* 349:293-299.)

20 Antibody fragments which contain specific binding sites for GCRP may also be generated. For example, such fragments include, but are not limited to, F(ab')2 fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) *Science* 246:1275-1281.)

25 Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between GCRP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two 30 non-interfering GCRP epitopes is preferred, but a competitive binding assay may also be employed (Pound, supra).

35 Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for GCRP. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of GCRP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined

for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple GCRP epitopes, represents the average affinity, or avidity, of the antibodies for GCRP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular GCRP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} 5 L/mole are preferred for use in immunoassays in which the GCRP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of GCRP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to 10 Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is preferred for use in procedures requiring precipitation of GCRP-antibody 15 complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding GCRP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the 20 polynucleotide encoding GCRP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding GCRP. Thus, complementary molecules or fragments may be used to modulate GCRP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along 25 the coding or control regions of sequences encoding GCRP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding GCRP. (See, 30 e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding GCRP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding GCRP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled 35 by endogenous nucleases. Transient expression may last for a month or more with a non-replicating

vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding GCRP. Oligonucleotides derived from the transcription initiation site, e.g.,

5 between about positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and
10 Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme
15 molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding GCRP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU,
20 and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

25 Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding GCRP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable
30 RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiester linkages within
35 the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in

all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for 5 use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) *Nat. Biotechnol.* 15:462-466.)

10 Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic 15 effects discussed above. Such pharmaceutical compositions may consist of GCRP, antibodies to GCRP, and mimetics, agonists, antagonists, or inhibitors of GCRP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in 20 combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

25 In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

30 Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

35 Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain

tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and 5 tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene 10 glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit 15 capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous 20 solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection 25 suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art. 30 The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acids. Salts 35 tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms.

In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate 5 container and labeled for treatment of an indicated condition. For administration of GCRP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

10 For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

15 A therapeutically effective dose refers to that amount of active ingredient, for example GCRP or fragments thereof, antibodies of GCRP, and agonists, antagonists or inhibitors of GCRP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the dose lethal to 50% of the 20 population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD₅₀/ED₅₀ ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED₅₀ with little or no toxicity. The dosage varies within this 25 range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of 30 the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μ g to 100,000 μ g, up to a total dose of about 1 35 gram, depending upon the route of administration. Guidance as to particular dosages and methods of

delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

5 In another embodiment, antibodies which specifically bind GCRP may be used for the diagnosis of disorders characterized by expression of GCRP, or in assays to monitor patients being treated with GCRP or agonists, antagonists, or inhibitors of GCRP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for GCRP include methods which utilize the antibody and a label to detect GCRP in human body fluids or in extracts of cells 10 or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring GCRP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of GCRP expression. Normal or 15 standard values for GCRP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to GCRP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, preferably by photometric means. Quantities of GCRP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject 20 values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding GCRP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of GCRP may be correlated with 25 disease. The diagnostic assay may be used to determine absence, presence, and excess expression of GCRP, and to monitor regulation of GCRP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding GCRP or closely related molecules may be used to identify nucleic acid sequences which encode GCRP. The specificity of the probe, whether it is made 30 from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low), will determine whether the probe identifies only naturally occurring sequences encoding GCRP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably have at 35 least 50% sequence identity to any of the GCRP encoding sequences. The hybridization probes of the

subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:9-16 or from genomic sequences including promoters, enhancers, and introns of the GCRP gene.

Means for producing specific hybridization probes for DNAs encoding GCRP include the cloning of polynucleotide sequences encoding GCRP or GCRP derivatives into vectors for the production of 5 mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ³²P or ³⁵S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

10 Polynucleotide sequences encoding GCRP may be used for the diagnosis of disorders associated with expression of GCRP. Examples of such disorders include, but are not limited to, a nervous disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, 15 retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, 20 cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorder of the central nervous system, cerebral palsy, a neuroskeletal disorder, an autonomic nervous system disorder, a cranial nerve disorder, a spinal cord disease, muscular dystrophy and other neuromuscular disorders, a peripheral nervous system disorder, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathy; myasthenia gravis, periodic paralysis; 25 a mental disorder including mood, anxiety, and schizophrenic disorders; seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary 30 thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; and an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), 35 Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis,

anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis,
5 glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner
10 syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma. The polynucleotide sequences encoding GCRP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered GCRP expression. Such qualitative or quantitative methods are
15 well known in the art.

In a particular aspect, the nucleotide sequences encoding GCRP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding GCRP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable
20 incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding GCRP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an
25 individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of GCRP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding GCRP, under conditions suitable for hybridization or amplification. Standard
30 hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

35 Once the presence of a disorder is established and a treatment protocol is initiated, hybridization

assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or 5 overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

10 Additional diagnostic uses for oligonucleotides designed from the sequences encoding GCRP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding GCRP, or a fragment of a polynucleotide complementary to the polynucleotide encoding GCRP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may 15 also be employed under less stringent conditions for detection or quantitation of closely related DNA or RNA sequences.

Methods which may also be used to quantify the expression of GCRP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) *J. Immunol. Methods* 159:235-244; Duplaa, C. et al. (1993) 20 *Anal. Biochem.* 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the 25 polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., 30 Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) *Proc. Natl. Acad. Sci. USA* 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding GCRP may be used to 35 generate hybridization probes useful in mapping the naturally occurring genomic sequence. The

sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355; Price,

5 C.M. (1993) *Blood Rev.* 7:127-134; and Trask, B.J. (1991) *Trends Genet.* 7:149-154.)

Fluorescent *in situ* hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, *supra*, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) site. Correlation between the location of the gene encoding

10 GCRP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

15 *In situ* hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery 20 techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) *Nature* 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or 25 affected individuals.

In another embodiment of the invention, GCRP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes 30 between GCRP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with GCRP, or fragments thereof, and washed. Bound 35 GCRP is then detected by methods well known in the art. Purified GCRP can also be coated directly onto

plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding GCRP specifically compete with a test compound for binding GCRP. In 5 this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with GCRP.

In additional embodiments, the nucleotide sequences which encode GCRP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on 10 properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific 15 embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. [Attorney Docket No. PF-0612 P, filed October 6, 1998] and U.S. Ser. No. 60/133,585, are hereby expressly incorporated by reference.

20

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic 25 solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo 30 d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA 35 libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP

vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the 5 appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), or pINCY (Incyte Pharmaceuticals, Palo Alto CA).

10 Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or 15 WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8, QIAWELL 8 Plus, QIAWELL 8 Ultra plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

20 Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) *Anal. Biochem.* 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy,

25 Helsinki, Finland).

III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the 30 MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 35 sequencing systems (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software;

or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed
5 using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth
10 column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal
15 algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a
20 selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the
25 corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.)

30 The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:9-
16. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

35 Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene

and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, *supra*, ch. 7; Ausubel, 1995, *supra*, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related 5 molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

10

100

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may 15 identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding GCRP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, 20 reproductive, and urologic. The disease/condition categories included cancer, inflammation/trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

V. Extension of GCRP Encoding Polynucleotides

25 The full length nucleic acid sequences of SEQ ID NO:9-16 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in 30 length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

35 High fidelity amplification was obtained by PCR using methods well known in the art. PCR was

performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg²⁺, (NH₄)₂SO₄, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI

5 A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

10 The concentration of DNA in each well was determined by dispensing 100 μl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of 15 DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun 20 sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, 25 individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 30 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulphoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction 35 kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:9-16 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

5 Hybridization probes derived from SEQ ID NO:9-16 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ - 32 P] adenosine 10 triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10⁷ counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA 15 digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate.

20 Hybridization patterns are visualized using autoradiography and compared.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, *supra*.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or 25 mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

30 Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a 35 glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and

chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

5 Sequences complementary to the GCRP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring GCRP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of GCRP. To inhibit transcription, a
10 complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the GCRP-encoding transcript.

IX. Expression of GCRP

Expression and purification of GCRP is achieved using bacterial or virus-based expression systems. For expression of GCRP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria
20 express GCRP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of GCRP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant *Autographica californica* nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding GCRP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates.
25 Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect *Spodoptera frugiperda* (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:3224-3227; Sandig, V. et al. (1996) *Hum. Gene Ther.* 7:1937-1945.)
30 In most expression systems, GCRP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from *Schistosoma japonicum*, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech).
35 Following purification, the GST moiety can be proteolytically cleaved from GCRP at specifically

engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10 and 16).

5 Purified GCRP obtained by these methods can be used directly in the following activity assay.

X. Demonstration of GCRP Activity

The activity of GCRP may be measured using an assay based upon the property of some GPCRs to support proliferation *in vitro* of fibroblasts and tumor cells under serum-free conditions (Chiquet-Ehrismann, R. et al. (1986) *Cell* 47:131-139). Wells in 96 well cluster plates (Falcon, Fisher Scientific, 10 Santa Clara CA) are coated with GCRP by incubation with solutions at 50-100 µg/ml for 15 min at ambient temperature. The coating solution is aspirated, and the wells washed with Dulbecco's medium before cells are plated. Rat fibroblast cultures or rat mammary tumor cells are prepared as described and plated at a density of 10⁴-10⁵ cells/ml in Dulbecco's medium supplemented with 10% fetal calf serum (FCS).
15 After three days the media are removed, and the cells washed three times with phosphate-buffered saline (PBS) before the addition of serum-free Dulbecco's medium containing 0.25 mg/ml bovine serum albumin (BSA, Fraction V, Sigma Chemical, St. Louis, MO). After 2 days the medium is aspirated, and 100 µl of [³H]thymidine (NEN) at 2 µCi/ml in fresh Dulbecco's medium containing 0.25 mg/ml BSA added. Parallel plates are fixed and stained to determine cell numbers. After 16 hr, the 20 medium is aspirated, the cell layer washed with PBS, and the 10% trichloroacetic acid-precipitable counts in the cell layer determined by liquid scintillation counting of radioisotope (normalized to relative cell numbers; Chiquet-Ehrismann, R. et al. (1986) *supra*). The rates of cell proliferation and [³H]thymidine uptake are proportional to the levels of GCRP in the sample.

In the alternative, the assay for GCRP activity is based upon the property of CD97/Emr1 GPCR 25 family proteins to modulate G protein-activated second messenger signal transduction pathways (e.g., cAMP; Gaudin, P. et al. (1998) *J. Biol. Chem.* 273:4990-4996). A plasmid encoding full length GCRP is transfected into a mammalian cell line (e.g., COS-7 or Chinese hamster ovary (CHO-K1) cell lines) using methods well-known in the art. Transfected cells are grown in 12-well trays in culture medium containing 2% FCS for 48 hours, the culture medium is discarded, then the attached cells are gently 30 washed with PBS. The cells are then incubated in culture medium with 10% FCS or 2% FCS for 30 minutes, then the medium is removed and cells lysed by treatment with 1 M perchloric acid. The cAMP levels in the lysate are measured by radioimmunoassay using methods well-known in the art. Changes in the levels of cAMP in the lysate from 10% FCS-treated cells compared with those in 2% FCS-treated cells are proportional to the amount of GCRP present in the transfected cells.

35 **XI. Functional Assays**

GCRP function is assessed by expressing the sequences encoding GCRP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which 5 contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker 10 proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as 15 measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in 20 flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of GCRP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding GCRP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads 25 coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding GCRP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XII. Production of GCRP Specific Antibodies

30 GCRP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the GCRP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is 35 synthesized and used to raise antibodies by means known to those of skill in the art. Methods for

selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) 5 by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide activity by, for example, binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

10 XIII. Purification of Naturally Occurring GCRP Using Specific Antibodies

Naturally occurring or recombinant GCRP is substantially purified by immunoaffinity chromatography using antibodies specific for GCRP. An immunoaffinity column is constructed by covalently coupling anti-GCRP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed 15 according to the manufacturer's instructions.

Media containing GCRP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of GCRP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/GCRP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotropic, such as urea or thiocyanate ion), and 20 GCRP is collected.

XIV. Identification of Molecules Which Interact with GCRP

GCRP, or biologically active fragments thereof, are labeled with ¹²⁵I Bolton-Hunter reagent. (See, e.g., Bolton et al. (1973) Biochem. J. 133:529.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled GCRP, washed, and any wells with labeled 25 GCRP complex are assayed. Data obtained using different concentrations of GCRP are used to calculate values for the number, affinity, and association of GCRP with the candidate molecules.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. 30 Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

| Polypeptide SEQ ID NO: | Nucleotide SEQ ID NO: | Clone ID | Library | Fragments | |
|---------------------------|--------------------------|-------------|-----------|---|--|
| | | | | | |
| 1 | 9 | 1595762 | BRAINOT14 | 1595762H1 (BRAINOT14), 1546740R6 (PROSTUT04), 1597820F6 (BRAINOT14), 2716791F3, (THYRNNOT09), 1705908F6 (DUODNOT02), 1753259F6 (LIVRTUT01), 1453415F1 (PENITUT01), 1314567F1 (BLADTUT02), 115604F1 (KIDNNNOT01), SBIA04259D1, SBIA03661D1, SBIA05044D1 | |
| 2 | 10 | 2763296 | BRSTNOT12 | 2763296H1 (BRSTNOT12), 3486025F6 (KIDNNNOT31), SBDAO7002F3 | |
| 3 | 11 | 3367641 | CONNUTU04 | 340333R6 (NEUTFMT01), 3367641F6 (CONNUTU04), 3367641H1 (CONNUTU04), 3367641T6 (CONNUTU04) | |
| 4 | 12 | 866390 | BRAITUT03 | 866390H1 (BRAITUT03), 866390T1 (BRAITUT03), 2112194R6 (BRAITUT03) | |
| 5 | 13 | 998550 | KIDNTUT01 | 998550H1 (KIDNTUT01), 998550X312D1 (KIDNTUT01), SXCA00752V1, SCDA06667V1, SXCA00292V1 | |
| 6 | 14 | 1547873 | PROSNOT06 | 980190R6 (TONGTUT01), 980190T1 (TONGTUT01), 1547873H1 (PROSNOT06), SBHA00047F1, SBHA00670F1 | |
| 7 | 15 | 2048484 | LIVRFET02 | 1872334F6 (LEUKNOT02), 1872334T3 (LEUKNOT02), 2048484H1 (LIVRFET02), SBGA01350F1, SBGA04183F1 | |
| 8 | 16 | 55544170 | TONSDIT01 | 1433664H1 (BEPINON01), 1522189F6 (BLADTUT04), 1542382R1 (SINTTUT01), 1662775T6 (BRSTNOT09), 1957981F6 (CONNNOT01), 3963055H1 (PROSNOT14), 5554170H1 (TONSDIT01), SASA02396F1, SASA00164F1 | |

Table 2

| SEQ ID NO: | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Genbank Homologs | Identification | Analytical Methods | |
|------------|---------------------|--|---|--|---------------------------------------|--------------------------------|--------------------------------|
| 1 | 693 | S201 T373 S80 S227 S253 S391 T469 S674 S36 T41 S113 S197 S218 T368 T442 T603 S660 S678 S690 Y558 | N39 N148 N171 N234 N303 N324 N341 | Signal Peptide: M1-G25 Transmembrane Domain: F578-I596, Y641-M661, L610-F630 Targeting Signal: S691 TM7XN1 protein [Homo sapiens] (g4456467) Seven transmembrane receptor HE6 [Homo sapiens] (g2117161) | Seven-span transmembrane receptor HE6 | | BLAST BLOCKS HMM MOTIFS PRINTS |
| 2 | 84 | S65 S36 T41 S51 S69 S81 | N39 | Signal Peptide: M1-G25 TM7XN1 protein [Homo sapiens] (g4456467) Seven transmembrane receptor HE6 [Homo sapiens] (g2117161) | Seven-span transmembrane receptor HE6 | BLAST HMM | |
| 3 | 136 | T80 T124 T64 S86 | N67 | Asx Hydroxylation site: C41-C52 EGF-Like Domain: P22-K37, C35-Y56 CD97 [Homo sapiens] (g3342740) | EGF-TM7 receptors family | BLAST BLOCKS HMM MOTIFS PRINTS | |

Table 2 (cont.)

| SEQ ID NO: | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Genbank Homologs | Identification | Analytical Methods |
|------------|---------------------|---------------------------------|-------------------------------|---|--------------------|--|
| 4 | 307 | S228 T48 S87 S93 T288 S297 | N5 | Transmembrane domain: I23-G41, L198-G41 Signal peptide: M1-S20 G-protein coupled receptors: G110-I126 Olfactory receptor: L59-R80, F177-D191, M237-G252, V271-L282, T288-F302 7-transmembrane receptor: G41-Y287 Olfactory receptor 4 [Gallus gallus] (g1246534) | Olfactory receptor | BLAST BLOCKS HMM MOTIFS PFAM PRINTS ProfileScan SPScan |
| 5 | 318 | S272 S57 S70 S111 S267 | N8 N45 N91 | Transmembrane domain: L34-V52, G202-L225 Signal peptide: M1-A26 G-protein coupled receptors: M113-I129 Olfactory receptor: M62-K83, S180-D194, F240-V255 7-transmembrane receptor: G44-Y294 Putative G-protein coupled receptor RA1c [Rattus norvegicus] (g3420759) | Olfactory receptor | BLAST BLOCKS HMM MOTIFS PFAM PRINTS ProfileScan SPScan |

Table 2 (cont.)

| SEQ ID NO: | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Genbank Homologs | Identification Methods |
|------------|---------------------|---------------------------------|-------------------------------|---|--|
| 6 | 138 | T6 S22 S125 | N44 N47 N83 | mas product homolog modulating intracellular angiotensin II [Homo sapiens] (g244210) | MAS-like GPCR BLAST |
| 7 | 199 | S21 S65 T93 S136 S155 T131 | | Transmembrane domain: L163-M184 Signal peptide: M1-G20 Rhodopsin-like GPCR superfamily: S160-M184 | Rhodopsin-like GPCR HMM PRINTS SPScan |
| 8 | 445 | S435 S273 | N129 | Transmembrane domain: V366-F386 Signal peptide: M1-A25 Contains similarity to G-coupled protein receptors [Caenorhabditis elegans] (g2736345) | GPCR BLAST HMM SPScan |

Table 3

| Nucleotide SEQ ID NO: | Range of Useful Nucleotide Fragment(s) | Tissue Expression (Fraction of Total) | Disease or Condition (Fraction of Total) | Vector |
|-----------------------|--|---|---|---------|
| 9 | 1879-1928 | Reproductive (0.290) Nervous (0.215) Gastrointestinal (0.131) | Cancer (0.495) Inflammation (0.168) Trauma (0.140) | pINCY |
| 10 | | Nervous (0.318) Reproductive (0.227) Urologic (0.136) | Cancer (0.636) Inflammation (0.136) Trauma (0.091) | pINCY |
| 11 | 325-387 667-732 | Hematopoietic/Immune (0.625) Reproductive (0.250) Musculoskeletal (0.125) | Cancer (0.500) Inflammation (0.400) | pINCY |
| 12 | 202-246 | Nervous (1.000) | Cancer and cell proliferation (1.000) | PSPORT1 |
| 13 | 226-270 | Reproductive (0.500) Urologic (0.500) | Cancer and cell proliferation (1.000) | PSPORT1 |
| 14 | 526-573 | Gastrointestinal (0.500) Reproductive (0.500) | Cancer and cell proliferation (1.000) | PSPORT1 |
| 15 | 191-235 | Developmental (0.167) Gastrointestinal (0.167) Hematopoietic/Immune (0.167) Musculoskeletal (0.167) Nervous (0.167) Reproductive (0.167) | Cancer and cell proliferation (0.834) Inflammation (0.167) | pINCY |
| 16 | 802-846 1627-1672 | Reproductive (0.290) Gastrointestinal (0.194) Nervous (0.161) | Cancer and cell proliferation (0.609) Inflammation (0.217) | pINCY |

Table 4

| Nucleotide SEQ ID NO: | Library | Library Comment |
|-----------------------|-----------|---|
| 9 | BRAINOT14 | Library was constructed using RNA isolated from brain tissue removed from the left frontal lobe of a 40-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology for the associated tumor tissue indicated grade 4 gemistocytic astrocytoma. |
| 10 | BRSTNOT12 | Library was constructed using RNA isolated from diseased breast tissue removed from a 32-year-old Caucasian female during a bilateral reduction mammoplasty. Pathology indicated nonproliferative fibrocystic disease. Family history included benign hypertension and atherosclerotic coronary artery disease. |
| 11 | CONNUT04 | Library was constructed using RNA isolated from tumorous spinal tissue removed from a 35-year-old Caucasian male during an exploratory laparotomy. Pathology indicated schwannoma with degenerative changes. Patient history included anxiety, depression, neurofibromatosis and benign neoplasm of the scrotum. Previously the patient had a spinal fusion. Family history included brain cancer, liver disease, and multiple sclerosis. |
| 12 | BRAITUT03 | Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe a 17-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology indicated a grade 4 fibrillary giant and small-cell astrocytoma. Family history included benign hypertension and cerebrovascular disease. |
| 13 | KIDNTUT01 | Library was constructed using RNA isolated from the kidney tumor tissue removed from an 8-month-old female during nephroureterectomy. Pathology indicated Wilms' tumor (nephroblastoma), which involved 90 percent of the renal parenchyma. Prior to surgery, the patient was receiving heparin anticoagulant therapy. |

Table 4 (cont.)

| Nucleotide SEQ ID NO: | Library | Library Comment |
|-----------------------|-----------|--|
| 14 | PROSNOT06 | Library was constructed using RNA isolated from the diseased prostate tissue of a 57-year-old Caucasian male during radical prostatectomy, removal of both testes and excision of regional lymph nodes. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated adenocarcinoma (Gleason grade 3+3). Patient history included a benign neoplasm of the large bowel and type I diabetes. Family history included a malignant neoplasm of the prostate and type I diabetes. |
| 15 | LIVRFET02 | Library was constructed using RNA isolated from liver tissue removed from a Caucasian female fetus, who died at 20 weeks' gestation. Family history included seven days of erythromycin treatment for bronchitis in the mother during the first trimester. |
| 16 | TONSDIT01 | Library was constructed using RNA isolated from the tonsil tissue of a 6-year-old Caucasian male during adenotonsillectomy. Pathology indicated lymphoid hyperplasia of the tonsils. The patient presented with an abscess of the pharynx. The patient was not taking any medications. Family history included hypothyroidism in the grandparent(s) and benign skin neoplasm in the sibling(s). |

Table 5

| Program | Description | Reference | Parameter Threshold |
|--------------------------|--|---|--|
| ABI FACTURA | A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences. | Perkin-Elmer Applied Biosystems, Foster City, CA. | |
| ABI/PARACEL FDF | A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences. | Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA. | Mismatch <50% |
| ABI AutoAssembler | A program that assembles nucleic acid sequences. | Perkin-Elmer Applied Biosystems, Foster City, CA. | |
| BLAST | A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx. | Altschul, S.F. et al. (1990) <i>J. Mol. Biol.</i> 215:403-410; Altschul, S.F. et al. (1997) <i>Nucleic Acids Res.</i> 25: 3389-3402. | <i>BLAST</i> : Probability value= 1.0E-8 or less <i>Assembled BLAST</i> : Probability value= 1.0E-10 or less <i>Full Length sequences</i> : Probability value= 1.0E-10 or less |
| FASTA | A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, usafa, fastx, usax, and ssearch. | Pearson, W.R. and D.J. Lipman (1988) <i>Proc. Natl. Acad. Sci.</i> 85:2444-2448; Pearson, W.R. (1990) <i>Methods Enzymol.</i> 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) <i>Adv. Appl. Math.</i> 2:482-489. | <i>BLAST</i> : Probability value= 1.0E-4, <i>Assembled BLAST</i> : fasta Identity= 95% or greater and Match length=200 bases or greater, fasta E value= 1.0E-8 or less <i>Full Length sequences</i> : fastx score=100 or greater |
| BLIMPS | A BLOCKs IMProved Searcher that matches a sequence against those in BLOCKS and PRINTS databases to search for gene families, sequence homology, and structural fingerprint regions. | Henikoff, S and J.G. Henikoff, <i>Nucl. Acid Res.</i> , 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) <i>Methods Enzymol.</i> 266:38-105; and Altwood, T.K. et al. (1997) <i>J. Chem. Inf. Comput. Sci.</i> 37: 417-424. | Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and Probability value = 1.0E-3 or less if applicable |
| PFAM | A Hidden Markov Models-based application useful for protein family search. | Krogh, A. et al. (1994) <i>J. Mol. Biol.</i> , 235:1501-1531; Sonnhammer, E.L.L. et al. (1998) <i>Nucleic Acids Res.</i> 26:320-322. | Score=10-50 bits, depending on individual protein families |

Table 5 cont.

| Program | Description | Reference | Parameter Threshold |
|-------------|---|--|--|
| ProfileScan | An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite. | Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Barroch, A. et al. (1997) Nucleic Acids Res. 25: 217-221. | Score= 4.0 or greater |
| Phred | A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability. | Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194. | |
| Phrap | A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences. | Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA. | Score= 120 or greater; Match length= 36 or greater |
| Consed | A graphical tool for viewing and editing Phrap assemblies | Gordon, D. et al. (1998) Genome Res. 8:195-202. | |
| SPScan | A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides. | Nelson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439. | Score=5 or greater |
| Motif | A program that searches amino acid sequences for patterns that matched those defined in Prosite. | Barroch et al. ^{suppl} ; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI. | |

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, and fragments thereof.
2. A substantially purified variant having at least 90% amino acid sequence identity to the amino acid sequence of claim 1.
- 10 3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.
4. An isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide of claim 3.
- 15 5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.
6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.
- 20 7. A method for detecting a polynucleotide, the method comprising the steps of:
 - (a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and
 - (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.
- 25 8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.
- 30 9. An isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, and fragments thereof.
- 35 10. An isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide of claim 9.

11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.

12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.

5

13. A host cell comprising the expression vector of claim 12.

14. A method for producing a polypeptide, the method comprising the steps of:

10 a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

15

16. A purified antibody which specifically binds to the polypeptide of claim 1.

17. A purified agonist of the polypeptide of claim 1.

20

18. A purified antagonist of the polypeptide of claim 1.

19. A method for treating or preventing a disorder associated with decreased expression or activity of GCRP, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.

25

20. A method for treating or preventing a disorder associated with increased expression or activity of GCRP, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

1/7

| | | | | | | | | | | | | | |
|-----|---|---|---|---|---|---|---|-------------------------------|---------|---|----------|------------|------------|
| 1 | M | - | - | - | - | - | - | T P Q S L L Q T T L F L L S L | - | - | [L F L V | 2763296 | |
| 1 | M | - | - | - | - | - | - | T P Q S L L Q T T L F L L S L | - | - | [L F L V | 1595762 | |
| 1 | M V F S V R Q C G H V G R T E E V L L T F K I F L L V I T C L H V V L V | | | | | | | | | | | GI 2117161 | |
| 21 | Q G A H G R G H R E D F R | - | - | - | - | - | - | F C S Q R N Q T H R S S | 2763296 | | | | |
| 21 | Q G A H G R G H R E D F R | - | - | - | - | - | - | F C S Q R N Q T H R S S | 1595762 | | | | |
| 36 | T S L E D T D N S S L S P P P A K L S V V S F A P S S N E V E T T S | | | | | | | | | | | GI 2117161 | |
| 46 | L H Y | - | - | - | - | - | - | - | - | - | - | | |
| 46 | L H Y K P T P D L R I S I E N S E E A L T V H A P F P A A H P A S R | - | - | - | - | - | - | - | - | - | - | 2763296 | |
| 71 | L N - D V T L S L L P S N E T E K T K I T I V K T F N A S G V K P Q R | | | | | | | | | | | GI 2117161 | |
| 49 | - | - | - | - | - | - | - | - | - | - | - | | |
| 80 | S F P D P R G L Y H F | - | - | - | - | - | - | C L Y W N R H | - | - | - | 1595762 | |
| 105 | N I C N L S S I C N D S A F F R G E I M F Q Y D K E S T V P Q N Q H I | | | | | | | | | | | I | GI 2117161 |
| 49 | - | - | - | - | - | - | - | - | - | - | - | | |
| 98 | A G R L H L L Y G K R D | - | - | - | - | - | - | [F L L] | - | - | - | 2763296 | |
| 140 | T N G T L T G V L S L S E L K R S E L N K T L Q T L S E T Y F I M C A | | | | | | | | | | | | |
| 49 | - | - | - | - | - | - | - | - | - | - | - | | |
| 113 | S D K A S S S L L C | - | - | - | - | - | - | - | - | - | - | 1595762 | |
| 175 | T A E A Q S T L N C T F T I K L N N T M N A C A A I A A L E R V K I R | | | | | | | | | | | GI 2117161 | |

FIGURE 1A

2/7

| | | | | | | | |
|----|-----|------------------------|------------------------|------------------|----------------------------|----------------------------|----------------|
| 49 | 122 | 131 | 149 | 155 | 179 | 199 | 209 |
| - | - | OGP | - | ASFTFSSH | LKRDLQLLSQFLKHPOKASRRPS | LAQSKLTS | LGSSLEPNL |
| - | - | DHP | - | ASPIGEIQPLSPQPS | MPPQSETTISSPMPPQTHVSGTPPPV | APPVQTDIVNTSISDLENQVLQMEKA | QVVSRLHSSPPDML |
| - | - | PPPPFSSQSI | - | PPHTAAHNASSVDMCE | KAASSPATAID | DRSFEEEDR | QVLAQPLAQ R |
| - | - | PMEHCCSVRI | PLLATSVTSWWSPQONISLPSA | PPHTAAHNASSVDMCE | GI 2117161 | 1595762 | GI 2117161 |
| - | - | PCPSSPEELGKLQCDLQDPIVC | PVVPRATVLSQVPKATSF | PPHTAAHNASSVDMCE | GI 2117161 | 1595762 | GI 2117161 |
| - | - | GI 2117161 | GI 2117161 | PPHTAAHNASSVDMCE | GI 2117161 | 1595762 | GI 2117161 |
| - | - | 2763296 | 2763296 | 2763296 | 2763296 | 2763296 | 2763296 |
| - | - | 1595762 | 1595762 | 1595762 | 1595762 | 1595762 | 1595762 |
| - | - | GI 2117161 | GI 2117161 | GI 2117161 | GI 2117161 | GI 2117161 | GI 2117161 |

FIGURE 1B

FIGURE 1C

FIGURE 1D

5/7

49 - - - - - **L G L S I L V L G L P W A L I F F S F A S G T F Q L V V L Y L F S** 2763296
 613 - - - - - **R S I A G L T F I L G I T W G F A F F A W - - G P V N V T F M Y L F A** 1595762
 830 GI 2117161

49 - - - - - **I I T S E Q G F F L I F I W Y - - - - - Y - - - - -**
 645 **I F N T I L Q G F F F I F I F Y C V A K E N V R K Q W R R Y L C C G K L R** 2763296
 863 GI 2117161

54 **L Q A R G G - - - - -**
 663 **L Q A R G G - - - - -**
 898 **L A E N S D W S K T A T N G L K K Q T V N Q G V S S S N S L Q S S S** 2763296
 GI 2117161

68 **D S A - - - - -**
 677 **D S A - - - - -**
 933 **N S T N S T T L L V N N D C S V H A S G N G N A S T E R N G V S F S V** 2763296
 GI 2117161

71 - - - - - **R L P I S S G 2763296**
 680 - - - - - **R L P I S S G 1595762**
 968 **Q N G D V C L H D F T G K Q H M F N E K E D S C N G K G R M A L R R T** 2763296
 GI 2117161

78 **S T S S S - - - - - R I**
 687 **S T S S S - - - - - R I**
 1003 **S K R G S L H F I E Q M** 1595762
 GI 2117161

FIGURE 1E

6/7

| | | | | | | |
|-----|---|---|-----------------|-----------------|-------------|------------|
| 1 | M G V S A L T E L T | - - - S L G G | - - - D S L S C | - - - D S R G C | A R W C | GI 3342740 |
| 20 | P Q D S S C V N A T A C R C N P G F S S F S E I I T T | - - - C V | - - - - - | - - - - - | P D I P D I | 3367641 |
| 31 | N E C A T P S K V S C G K F S D C W N T E G S Y D C V C S P | N E C A T L S K V S C G K F S D C W N T E G S Y D C V C S P | 3367641 | 3367641 | GI 3342740 | GI 3342740 |
| 55 | G Y E P V S G A K T F K N E S E N T C Q G K N H P T S S D F | G Y E P V S G A K T F K N E S E N T C Q G K N H P T S S D F | 3367641 | 3367641 | GI 3342740 | GI 3342740 |
| 91 | P S M R F G V T R A I L A A S R E Q G L G | P S M R F G V T R A I L A A S R E Q G L G | - - - - - | - - - - - | W D R - | 3367641 |
| 85 | L C K S Y G T C V N T L G S Y T C Q C L P G F K D V N E C T | L C K S Y G T C V N T L G S Y T C Q C L P G F K D V N E C T | 3367641 | 3367641 | GI 3342740 | GI 3342740 |
| 121 | S G Q N P C H S S T H C L N N V G S Y Q C R C R P G W Q P I | S G Q N P C H S S T H C L N N V G S Y Q C R C R P G W Q P I | 3367641 | 3367641 | GI 3342740 | GI 3342740 |
| 106 | - - - C R V S A W R L R | - - - C R V S A W R L R | - - - - - | - - - - - | W D R - | 3367641 |
| 151 | - - - S G Q N P C H S S T H C L N N V G S Y Q C R C R P G W Q P I | - - - S G Q N P C H S S T H C L N N V G S Y Q C R C R P G W Q P I | 3367641 | 3367641 | GI 3342740 | GI 3342740 |

FIGURE 2A

7/7

118 - - - - -
181 P G S P N G P N N T V C E D V D E C S S G Q H Q C D S S T V GI 3342740
118 - - - - -
211 C F N T V G S Y S C R C R P G W K P R H G I P N N Q K D T V GI 3367641
121 - - - - **T Y P T** - - - **T P E R Q S D V** - - - - -
241 C E D M **T F S T** W T P P **P G V H S Q T L** S R F F D K V Q D L GI 3367641
133 - - - - -
271 G R D H L S S F A V L M A H Y D V
3367641
GI 3342740

FIGURE 2B

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.

TANG, Y. Tom
YUE, Henry
LAL, Preeti
BANDMAN, Olga
AU-YOUNG, Janice
REDDY, Roopa
CORLEY, Neil C.
GUEGLER, Karl J.
GORONE, Gina A.
BAUGHN, Mariah R.
AZIMZAI, Yalda

<120> G-PROTEIN COUPLED RECEPTOR PROTEINS

<130> PF-0612 PCT

<140> To Be Assigned
<141> Herewith

<150> 09/167,219; unassigned; 60/133,585
<151> 1998-10-06; 1998-10-06; 1999-05-11

<160> 18

<170> PERL Program

<210> 1
<211> 693
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1595762CD1

<400> 1

Met Thr Pro Gln Ser Leu Leu Gln Thr Thr Leu Phe Leu Leu Ser
1 5 10 15
Leu Leu Phe Leu Val Gln Gly Ala His Gly Arg Gly His Arg Glu
20 25 30
Asp Phe Arg Phe Cys Ser Gln Arg Asn Gln Thr His Arg Ser Ser
35 40 45
Leu His Tyr Lys Pro Thr Pro Asp Leu Arg Ile Ser Ile Glu Asn
50 55 60
Ser Glu Glu Ala Leu Thr Val His Ala Pro Phe Pro Ala Ala His
65 70 75
Pro Ala Ser Arg Ser Phe Pro Asp Pro Arg Gly Leu Tyr His Phe
80 85 90
Cys Leu Tyr Trp Asn Arg His Ala Gly Arg Leu His Leu Leu Tyr
95 100 105
Gly Lys Arg Asp Phe Leu Leu Ser Asp Lys Ala Ser Ser Leu Leu

| | | |
|-------------------------------------|-------------------------|-----|
| 110 | 115 | 120 |
| Cys Phe Gln His Gln Glu Glu Ser Leu | Ala Gln Gly Pro Pro Leu | |
| 125 | 130 | 135 |
| Leu Ala Thr Ser Val Thr Ser Trp Trp | Ser Pro Gln Asn Ile Ser | |
| 140 | 145 | 150 |
| Leu Pro Ser Ala Ala Ser Phe Thr Phe | Ser Phe His Ser Pro Pro | |
| 155 | 160 | 165 |
| His Thr Ala Ala His Asn Ala Ser Val | Asp Met Cys Glu Leu Lys | |
| 170 | 175 | 180 |
| Arg Asp Leu Gln Leu Leu Ser Gln Phe | Leu Lys His Pro Gln Lys | |
| 185 | 190 | 195 |
| Ala Ser Arg Arg Pro Ser Ala Ala Pro | Ala Ser Gln Gln Leu Gln | |
| 200 | 205 | 210 |
| Ser Leu Glu Ser Lys Leu Thr Ser Val | Arg Phe Met Gly Asp Met | |
| 215 | 220 | 225 |
| Val Ser Phe Glu Glu Asp Arg Ile Asn | Ala Thr Val Trp Lys Leu | |
| 230 | 235 | 240 |
| Gln Pro Thr Ala Gly Leu Gln Asp Leu | His Ile His Ser Arg Gln | |
| 245 | 250 | 255 |
| Glu Glu Glu Gln Ser Glu Ile Met Glu | Tyr Ser Val Leu Leu Pro | |
| 260 | 265 | 270 |
| Arg Thr Leu Phe Gln Arg Thr Lys Gly | Arg Ser Gly Glu Ala Glu | |
| 275 | 280 | 285 |
| Lys Arg Leu Leu Leu Val Asp Phe Ser | Ser Gln Ala Leu Phe Gln | |
| 290 | 295 | 300 |
| Asp Lys Asn Ser Ser Gln Val Leu Gly | Glu Lys Val Leu Gly Ile | |
| 305 | 310 | 315 |
| Val Val Gln Asn Thr Lys Val Ala Asn | Leu Thr Glu Pro Val Val | |
| 320 | 325 | 330 |
| Leu Thr Phe Gln His Gln Leu Gln Pro | Lys Asn Val Thr Leu Gln | |
| 335 | 340 | 345 |
| Cys Val Phe Trp Val Glu Asp Pro Thr | Leu Ser Ser Pro Gly His | |
| 350 | 355 | 360 |
| Trp Ser Ser Ala Gly Cys Glu Thr Val | Arg Arg Glu Thr Gln Thr | |
| 365 | 370 | 375 |
| Ser Cys Phe Cys Asn His Leu Thr Tyr | Phe Ala Val Leu Met Val | |
| 380 | 385 | 390 |
| Ser Ser Val Glu Val Asp Ala Val His | Lys His Tyr Leu Ser Leu | |
| 395 | 400 | 405 |
| Leu Ser Tyr Val Gly Cys Val Val Ser | Ala Leu Ala Cys Leu Val | |
| 410 | 415 | 420 |
| Thr Ile Ala Ala Tyr Leu Cys Ser Arg | Val Pro Leu Pro Cys Arg | |
| 425 | 430 | 435 |
| Arg Lys Pro Arg Asp Tyr Thr Ile Lys | Val His Met Asn Leu Leu | |
| 440 | 445 | 450 |
| Leu Ala Val Phe Leu Leu Asp Thr Ser | Phe Leu Leu Ser Glu Pro | |
| 455 | 460 | 465 |
| Val Ala Leu Thr Gly Ser Glu Ala Gly | Cys Arg Ala Ser Ala Ile | |
| 470 | 475 | 480 |
| Phe Leu His Phe Ser Leu Leu Thr Cys | Leu Ser Trp Met Gly Leu | |
| 485 | 490 | 495 |
| Glu Gly Tyr Asn Leu Tyr Arg Leu Val | Val Glu Val Phe Gly Thr | |
| 500 | 505 | 510 |
| Tyr Val Pro Gly Tyr Leu Leu Lys Leu | Ser Ala Met Gly Trp Gly | |
| 515 | 520 | 525 |

Phe Pro Ile Phe Leu Val Thr Leu Val Ala Leu Val Asp Val Asp
 530 535 540
 Asn Tyr Gly Pro Ile Ile Leu Ala Val His Arg Thr Pro Glu Gly
 545 550 555
 Val Ile Tyr Pro Ser Met Cys Trp Ile Arg Asp Ser Leu Val Ser
 560 565 570
 Tyr Ile Thr Asn Leu Gly Leu Phe Ser Leu Val Phe Leu Phe Asn
 575 580 585
 Met Ala Met Leu Ala Thr Met Val Val Gln Ile Leu Arg Leu Arg
 590 595 600
 Pro His Thr Gln Lys Trp Ser His Val Leu Thr Leu Leu Gly Leu
 605 610 615
 Ser Leu Val Leu Gly Leu Pro Trp Ala Leu Ile Phe Phe Ser Phe
 620 625 630
 Ala Ser Gly Thr Phe Gln Leu Val Val Leu Tyr Leu Phe Ser Ile
 635 640 645
 Ile Thr Ser Phe Gln Gly Phe Leu Ile Phe Ile Trp Tyr Trp Ser
 650 655 660
 Met Arg Leu Gln Ala Arg Gly Gly Pro Ser Pro Leu Lys Ser Asn
 665 670 675
 Ser Asp Ser Ala Arg Leu Pro Ile Ser Ser Gly Ser Thr Ser Ser
 680 685 690
 Ser Arg Ile

<210> 2
 <211> 84
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2763296CD1

<400> 2
 Met Thr Pro Gln Ser Leu Leu Gln Thr Thr Leu Phe Leu Leu Ser
 1 5 10 15
 Leu Leu Phe Leu Val Gln Gly Ala His Gly Arg Gly His Arg Glu
 20 25 30
 Asp Phe Arg Phe Cys Ser Gln Arg Asn Gln Thr His Arg Ser Ser
 35 40 45
 Leu His Tyr Tyr Trp Ser Met Arg Leu Gln Ala Arg Gly Gly Pro
 50 55 60
 Ser Pro Leu Lys Ser Asn Ser Asp Ser Ala Arg Leu Pro Ile Ser
 65 70 75
 Ser Gly Ser Thr Ser Ser Arg Ile
 80

<210> 3
 <211> 136
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3367641CD1

<400> 3

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Gly | Val | Ser | Ala | Leu | Thr | Glu | Leu | Thr | Ser | Leu | Gly | Gly | Asp |
| 1 | | 5 | | | | 10 | | | | | | | | 15 |
| Ser | Leu | Ser | Cys | Cys | Val | Pro | Asp | Ile | Asn | Glu | Cys | Ala | Thr | Pro |
| | | | | | 20 | | | | 25 | | | | | 30 |
| Ser | Lys | Val | Ser | Cys | Gly | Lys | Phe | Ser | Asp | Cys | Trp | Asn | Thr | Glu |
| | | | | | 35 | | | | 40 | | | | | 45 |
| Gly | Ser | Tyr | Asp | Cys | Val | Cys | Ser | Pro | Gly | Tyr | Glu | Pro | Val | Ser |
| | | | | | 50 | | | | 55 | | | | | 60 |
| Gly | Ala | Lys | Thr | Phe | Lys | Asn | Glu | Ser | Glu | Asn | Thr | Cys | Gln | Gly |
| | | | | | 65 | | | | 70 | | | | | 75 |
| Lys | Asn | His | Pro | Thr | Ser | Ser | Asp | Phe | Pro | Ser | Met | Arg | Phe | Gly |
| | | | | | 80 | | | | 85 | | | | | 90 |
| Val | Thr | Arg | Ala | Ile | Leu | Ala | Ala | Ser | Arg | Glu | Gln | Gly | Leu | Gly |
| | | | | | 95 | | | | 100 | | | | | 105 |
| Cys | Arg | Val | Ser | Ala | Trp | Arg | Leu | Arg | Trp | Asp | Arg | Cys | Thr | Cys |
| | | | | | 110 | | | | 115 | | | | | 120 |
| Thr | Tyr | Pro | Thr | Thr | Pro | Glu | Arg | Gln | Ser | Asp | Val | Arg | Ile | Lys |
| | | | | | 125 | | | | 130 | | | | | 135 |
| Gly | | | | | | | | | | | | | | |

<210> 4
 <211> 307
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 866390CD1

<400> 4

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Glu | Ser | Glu | Asn | Arg | Thr | Val | Ile | Arg | Glu | Phe | Ile | Leu | Leu |
| 1 | | | | 5 | | | | 10 | | | | | | 15 |
| Gly | Leu | Thr | Gln | Ser | Gln | Asp | Ile | Gln | Leu | Leu | Val | Phe | Val | Leu |
| | | | | | 20 | | | | 25 | | | | | 30 |
| Val | Leu | Ile | Phe | Tyr | Phe | Ile | Ile | Leu | Pro | Gly | Asn | Phe | Leu | Ile |
| | | | | | 35 | | | | 40 | | | | | 45 |
| Ile | Phe | Thr | Ile | Lys | Ser | Asp | Pro | Gly | Leu | Thr | Ala | Pro | Leu | Tyr |
| | | | | | 50 | | | | 55 | | | | | 60 |
| Phe | Phe | Leu | Gly | Asn | Leu | Ala | Phe | Leu | Asp | Ala | Ser | Tyr | Ser | Phe |
| | | | | | 65 | | | | 70 | | | | | 75 |
| Ile | Val | Ala | Pro | Arg | Met | Leu | Val | Asp | Phe | Leu | Ser | Ala | Lys | Lys |
| | | | | | 80 | | | | 85 | | | | | 90 |
| Ile | Ile | Ser | Tyr | Arg | Gly | Cys | Ile | Thr | Gln | Leu | Phe | Phe | Leu | His |
| | | | | | 95 | | | | 100 | | | | | 105 |
| Phe | Leu | Gly | Gly | Glu | Gly | Leu | Leu | Leu | Val | Val | Met | Ala | Phe | |
| | | | | | 110 | | | | 115 | | | | | 120 |
| Asp | Arg | Tyr | Ile | Ala | Ile | Cys | Arg | Pro | Leu | His | Tyr | Pro | Thr | Val |
| | | | | | 125 | | | | 130 | | | | | 135 |
| Met | Asn | Pro | Arg | Thr | Cys | Tyr | Ala | Met | Met | Leu | Ala | Leu | Trp | Leu |

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 140 | 145 | 150 | | | | | | | | | | | | |
| Gly | Gly | Phe | Val | His | Ser | Ile | Ile | Gln | Val | Val | Leu | Ile | Leu | Arg |
| 155 | 160 | 165 | | | | | | | | | | | | |
| Leu | Pro | Phe | Cys | Gly | Pro | Asn | Gln | Leu | Asp | Asn | Phe | Phe | Cys | Asp |
| 170 | 175 | 180 | | | | | | | | | | | | |
| Val | Pro | Gln | Val | Ile | Lys | Leu | Ala | Cys | Thr | Asp | Thr | Phe | Val | Val |
| 185 | 190 | 195 | | | | | | | | | | | | |
| Glu | Leu | Leu | Met | Val | Phe | Asn | Ser | Gly | Leu | Met | Thr | Leu | Leu | Cys |
| 200 | 205 | 210 | | | | | | | | | | | | |
| Phe | Leu | Gly | Leu | Leu | Ala | Ser | Tyr | Ala | Val | Ile | Leu | Cys | Arg | Ile |
| 215 | 220 | 225 | | | | | | | | | | | | |
| Arg | Gly | Ser | Ser | Ser | Glu | Ala | Lys | Asn | Lys | Ala | Met | Ser | Thr | Cys |
| 230 | 235 | 240 | | | | | | | | | | | | |
| Ile | Thr | His | Ile | Ile | Val | Ile | Phe | Phe | Met | Phe | Gly | Pro | Gly | Ile |
| 245 | 250 | 255 | | | | | | | | | | | | |
| Phe | Ile | Tyr | Thr | Arg | Pro | Phe | Arg | Ala | Phe | Pro | Ala | Asp | Lys | Val |
| 260 | 265 | 270 | | | | | | | | | | | | |
| Val | Ser | Leu | Phe | His | Thr | Val | Ile | Phe | Pro | Leu | Leu | Asn | Pro | Val |
| 275 | 280 | 285 | | | | | | | | | | | | |
| Ile | Tyr | Thr | Leu | Arg | Asn | Gln | Glu | Val | Lys | Ala | Ser | Met | Lys | Lys |
| 290 | 295 | 300 | | | | | | | | | | | | |
| Val | Phe | Asn | Lys | His | Ile | Ala | | | | | | | | |
| | 305 | | | | | | | | | | | | | |

<210> 5
 <211> 318
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 998550CD1

| | | | | | | | | | | | | | | |
|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| <400> 5 | | | | | | | | | | | | | | |
| Met | Met | Val | Asp | Pro | Asn | Gly | Asn | Glu | Ser | Ser | Ala | Thr | Tyr | Phe |
| 1 | 5 | 10 | 15 | | | | | | | | | | | |
| Ile | Leu | Ile | Gly | Leu | Pro | Gly | Leu | Glu | Glu | Ala | Gln | Phe | Trp | Leu |
| 20 | 25 | 30 | | | | | | | | | | | | |
| Ala | Phe | Pro | Leu | Cys | Ser | Leu | Tyr | Leu | Ile | Ala | Val | Leu | Gly | Asn |
| 35 | 40 | 45 | | | | | | | | | | | | |
| Leu | Thr | Ile | Ile | Tyr | Ile | Val | Arg | Thr | Glu | His | Ser | Leu | His | Glu |
| 50 | 55 | 60 | | | | | | | | | | | | |
| Pro | Met | Tyr | Ile | Phe | Leu | Cys | Met | Leu | Ser | Gly | Ile | Asp | Ile | Leu |
| 65 | 70 | 75 | | | | | | | | | | | | |
| Ile | Ser | Thr | Ser | Ser | Met | Pro | Lys | Met | Leu | Ala | Ile | Phe | Trp | Phe |
| 80 | 85 | 90 | | | | | | | | | | | | |
| Asn | Ser | Thr | Thr | Ile | Gln | Phe | Asp | Ala | Cys | Leu | Leu | Gln | Met | Phe |
| 95 | 100 | 105 | | | | | | | | | | | | |
| Ala | Ile | His | Ser | Leu | Ser | Gly | Met | Glu | Ser | Thr | Val | Leu | Leu | Ala |
| 110 | 115 | 120 | | | | | | | | | | | | |
| Met | Ala | Phe | Asp | Arg | Tyr | Val | Ala | Ile | Cys | His | Pro | Leu | Arg | His |
| 125 | 130 | 135 | | | | | | | | | | | | |
| Ala | Thr | Val | Leu | Thr | Leu | Pro | Arg | Val | Thr | Lys | Ile | Gly | Val | Ala |
| 140 | 145 | 150 | | | | | | | | | | | | |

Ala Val Val Arg Gly Ala Ala Leu Met Ala Pro Leu Pro Val Phe
 155 160 165
 Ile Lys Gln Leu Pro Phe Cys Arg Ser Asn Ile Leu Ser His Ser
 170 175 180
 Tyr Cys Leu His Gln Asp Val Met Lys Leu Ala Cys Asp Asp Ile
 185 190 195
 Arg Val Asn Val Val Tyr Gly Leu Ile Val Ile Ile Ser Ala Ile
 200 205 210
 Gly Leu Asp Ser Leu Leu Ile Ser Phe Ser Tyr Leu Leu Ile Leu
 215 220 225
 Lys Thr Val Leu Gly Leu Thr Arg Glu Ala Gln Ala Lys Ala Phe
 230 235 240
 Gly Thr Cys Val Ser His Val Cys Ala Val Phe Ile Phe Tyr Val
 245 250 255
 Pro Phe Ile Gly Leu Ser Met Val His Arg Phe Ser Lys Arg Arg
 260 265 270
 Asp Ser Pro Leu Pro Val Ile Leu Ala Asn Ile Tyr Leu Leu Val
 275 280 285
 Pro Pro Val Leu Asn Pro Ile Val Tyr Gly Val Lys Thr Lys Glu
 290 295 300
 Ile Arg Gln Arg Ile Leu Arg Leu Phe His Val Ala Thr His Ala
 305 310 315
 Ser Glu Pro

<210> 6
 <211> 138
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1547873CD1

<400> 6
 Met Met Pro Thr Gln Thr Leu Pro Glu Thr His Ala Pro Gly Glu
 1 5 10 15
 Thr Gln Ile Ser Leu Phe Ser Ser Gly Asp Gln Glu Ala Gln Asn
 20 25 30
 Pro Asn Leu Val Ser Gln Leu Cys Gly Val Phe Leu Gln Asn Glu
 35 40 45
 Thr Asn Glu Thr Ile His Met Gln Met Ser Met Ala Val Gly Gln
 50 55 60
 Gln Ala Leu Pro Leu Asn Ile Ile Ala Pro Lys Ala Val Leu Val
 65 70 75
 Ser Leu Cys Gly Val Leu Leu Asn Gly Thr Val Phe Trp Leu Leu
 80 85 90
 Cys Cys Gly Ala Thr Asn Pro Tyr Met Val Tyr Ile Leu His Leu
 95 100 105
 Val Ala Ala Asp Val Ile Tyr Leu Cys Cys Ser Ala Val Gly Phe
 110 115 120
 Leu Gln Ile Tyr Ser Val Ala Glu Asp Met Val Ser Phe Val His
 125 130 135
 His Gly Val

<210> 7
<211> 199
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2048484CD1

<400> 7

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Gly | Leu | Thr | Leu | Leu | Leu | Leu | Leu | Leu | Gly | Leu | Glu | Gly | |
| 1 | | | | | | | | | | 10 | | 15 | | |
| Gln | Gly | Ile | Val | Gly | Ser | Leu | Pro | Glu | Val | Leu | Gln | Ala | Pro | Val |
| | | | | | | | | | | 20 | | 25 | | 30 |
| Gly | Ser | Ser | Ile | Leu | Val | Gln | Cys | His | Tyr | Arg | Leu | Gln | Asp | Val |
| | | | | | | | | | | 35 | | 40 | | 45 |
| Lys | Ala | Gln | Lys | Val | Trp | Cys | Arg | Phe | Leu | Pro | Glu | Gly | Cys | Gln |
| | | | | | | | | | | 50 | | 55 | | 60 |
| Pro | Leu | Val | Ser | Ser | Ala | Val | Asp | Arg | Arg | Ala | Pro | Ala | Gly | Arg |
| | | | | | | | | | | 65 | | 70 | | 75 |
| Arg | Thr | Phe | Leu | Thr | Asp | Leu | Gly | Gly | Gly | Leu | Leu | Gln | Val | Glu |
| | | | | | | | | | | 80 | | 85 | | 90 |
| Met | Val | Thr | Leu | Gln | Glu | Glu | Asp | Ala | Gly | Glu | Tyr | Gly | Cys | Met |
| | | | | | | | | | | 95 | | 100 | | 105 |
| Val | Asp | Gly | Ala | Arg | Gly | Pro | Gln | Ile | Leu | His | Arg | Val | Ser | Leu |
| | | | | | | | | | | 110 | | 115 | | 120 |
| Asn | Ile | Leu | Pro | Pro | Glu | Glu | Glu | Glu | Thr | His | Lys | Ile | Gly | |
| | | | | | | | | | | 125 | | 130 | | 135 |
| Ser | Leu | Ala | Glu | Asn | Ala | Phe | Ser | Asp | Pro | Ala | Gly | Ser | Ala | Asn |
| | | | | | | | | | | 140 | | 145 | | 150 |
| Pro | Leu | Glu | Pro | Ser | Gln | Asp | Glu | Lys | Ser | Ile | Pro | Leu | Ile | Trp |
| | | | | | | | | | | 155 | | 160 | | 165 |
| Gly | Ala | Val | Leu | Leu | Val | Gly | Leu | Leu | Val | Ala | Ala | Val | Val | Leu |
| | | | | | | | | | | 170 | | 175 | | 180 |
| Phe | Ala | Val | Met | Ala | Lys | Arg | Lys | Gln | Glu | Ser | Leu | Leu | Ser | Gly |
| | | | | | | | | | | 185 | | 190 | | 195 |
| Pro | Pro | Arg | Gln | | | | | | | | | | | |

<210> 8
<211> 445
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5554170CD1

<400> 8

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ala | Ala | Pro | Thr | Pro | Ala | Arg | Pro | Val | Leu | Thr | His | Leu | Leu |
| 1 | | | | | | | | | | 5 | | 10 | | 15 |
| Val | Ala | Leu | Phe | Gly | Met | Gly | Ser | Trp | Ala | Ala | Val | Asn | Gly | Ile |
| | | | | | | | | | | 20 | | 25 | | 30 |

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Trp | Val | Glu | Leu | Pro | Val | Val | Val | Lys | Glu | Leu | Pro | Glu | Gly | Trp |
| | | | | 35 | | | | | 40 | | | | | 45 |
| Ser | Leu | Pro | Ser | Tyr | Val | Ser | Val | Leu | Val | Ala | Leu | Gly | Asn | Leu |
| | | | | 50 | | | | | 55 | | | | | 60 |
| Gly | Leu | Leu | Val | Val | Thr | Leu | Trp | Arg | Arg | Leu | Ala | Pro | Gly | Lys |
| | | | | 65 | | | | | 70 | | | | | 75 |
| Asp | Glu | Gln | Val | Pro | Ile | Arg | Val | Val | Gln | Val | Leu | Gly | Met | Val |
| | | | | 80 | | | | | 85 | | | | | 90 |
| Gly | Thr | Ala | Leu | Leu | Ala | Ser | Leu | Trp | His | His | Val | Ala | Pro | Val |
| | | | | 95 | | | | | 100 | | | | | 105 |
| Ala | Gly | Gln | Leu | His | Ser | Val | Ala | Phe | Leu | Ala | Leu | Ala | Phe | Val |
| | | | | 110 | | | | | 115 | | | | | 120 |
| Leu | Ala | Leu | Ala | Cys | Cys | Ala | Ser | Asn | Val | Thr | Phe | Leu | Pro | Phe |
| | | | | 125 | | | | | 130 | | | | | 135 |
| Leu | Ser | His | Leu | Pro | Pro | Arg | Phe | Leu | Arg | Ser | Phe | Phe | Leu | Gly |
| | | | | 140 | | | | | 145 | | | | | 150 |
| Gln | Gly | Leu | Ser | Ala | Leu | Leu | Pro | Cys | Val | Leu | Ala | Leu | Val | Gln |
| | | | | 155 | | | | | 160 | | | | | 165 |
| Gly | Val | Gly | Arg | Leu | Glu | Cys | Pro | Pro | Ala | Pro | Ile | Asn | Gly | Thr |
| | | | | 170 | | | | | 175 | | | | | 180 |
| Pro | Gly | Pro | Pro | Leu | Asp | Phe | Leu | Glu | Arg | Phe | Pro | Ala | Ser | Thr |
| | | | | 185 | | | | | 190 | | | | | 195 |
| Phe | Phe | Trp | Ala | Leu | Thr | Ala | Leu | Leu | Val | Ala | Ser | Ala | Ala | Ala |
| | | | | 200 | | | | | 205 | | | | | 210 |
| Phe | Gln | Gly | Leu | Leu | Leu | Leu | Pro | Pro | Pro | Pro | Ser | Val | Pro | Pro |
| | | | | 215 | | | | | 220 | | | | | 225 |
| Thr | Gly | Glu | Leu | Gly | Ser | Gly | Leu | Gln | Val | Gly | Ala | Pro | Gly | Ala |
| | | | | 230 | | | | | 235 | | | | | 240 |
| Glu | Glu | Glu | Val | Glu | Glu | Ser | Ser | Pro | Leu | Gln | Glu | Pro | Pro | Ser |
| | | | | 245 | | | | | 250 | | | | | 255 |
| Gln | Ala | Ala | Gly | Thr | Thr | Pro | Gly | Pro | Asp | Pro | Lys | Ala | Tyr | Gln |
| | | | | 260 | | | | | 265 | | | | | 270 |
| Leu | Leu | Ser | Ala | Arg | Ser | Ala | Cys | Leu | Leu | Gly | Leu | Leu | Ala | Ala |
| | | | | 275 | | | | | 280 | | | | | 285 |
| Thr | Asn | Ala | Leu | Thr | Asn | Gly | Val | Leu | Pro | Ala | Val | Gln | Ser | Phe |
| | | | | 290 | | | | | 295 | | | | | 300 |
| Ser | Cys | Leu | Pro | Tyr | Gly | Arg | Leu | Ala | Tyr | His | Leu | Ala | Val | Val |
| | | | | 305 | | | | | 310 | | | | | 315 |
| Leu | Gly | Ser | Ala | Ala | Asn | Pro | Leu | Ala | Cys | Phe | Leu | Ala | Met | Gly |
| | | | | 320 | | | | | 325 | | | | | 330 |
| Val | Leu | Cys | Arg | Ser | Leu | Ala | Gly | Leu | Gly | Gly | Leu | Ser | Leu | Leu |
| | | | | 335 | | | | | 340 | | | | | 345 |
| Gly | Val | Phe | Cys | Gly | Gly | Tyr | Leu | Met | Ala | Leu | Ala | Val | Leu | Ser |
| | | | | 350 | | | | | 355 | | | | | 360 |
| Pro | Cys | Pro | Pro | Leu | Val | Gly | Thr | Ser | Ala | Gly | Val | Val | Leu | Val |
| | | | | 365 | | | | | 370 | | | | | 375 |
| Val | Leu | Ser | Trp | Val | Leu | Cys | Leu | Gly | Val | Phe | Ser | Tyr | Val | Lys |
| | | | | 380 | | | | | 385 | | | | | 390 |
| Val | Ala | Ala | Ser | Ser | Leu | Leu | His | Gly | Gly | Gly | Arg | Pro | Ala | Leu |
| | | | | 395 | | | | | 400 | | | | | 405 |
| Leu | Ala | Ala | Gly | Val | Ala | Ile | Gln | Val | Gly | Ser | Leu | Leu | Gly | Ala |
| | | | | 410 | | | | | 415 | | | | | 420 |
| Val | Ala | Met | Phe | Pro | Pro | Thr | Ser | Ile | Tyr | His | Val | Phe | His | Ser |
| | | | | 425 | | | | | 430 | | | | | 435 |

Arg Lys Asp Cys Ala Asp Pro Cys Asp Ser
440 445

<210> 9
<211> 3632
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1595762CB1

<400> 9
ggaagggag gagcaggcca cacaggcaca ggccggtag ggacctgcc agacctggag 60
gtggtaact ccaagagtga ctccgtcgga ggaaaatgac tccccagtc ctgctgcaga 120
cgacactgtt cctgctgagt ctgctttcc tggccaagg tgcccacggc agggccaca 180
ggaaagactt tcgttctgc agccagcga accagacaca caggagcgc ctccactaca 240
aaccacacc agacctgcgc atctccatcg agaactccga agaggccctc acagtccatg 300
ccccttccc tgcagccac cctgttccc gatccccc tgaccccagg ggccttacc 360
acttctgcct ctactggaac cgacatgtg ggagattaca tcttcttat ggcaagcgtg 420
acttcttgcg gatgtgacaaa gcctctagcc tcctctgtt ccagcaccag gaggagagcc 480
tggctcaggc ccccccgtg ttagccactt ctgtcaccc tcgttggcgc cctcagaaca 540
tcagcctgcc cagtggccg agcttcaact tctcccca cagtcctccc cacacggccg 600
ctcacaatgc ctgggtggac atgtgcgagc tcaaaaggga cctccagtcg ctcagccagt 660
tcctgaagca tccccagaag gcctcaagga ggccctcgcc tgccccggcc agccagcagt 720
tgcagagcct ggagtgcggg ctgacccctg tgagattcat gggggacatg gtgtcttcg 780
aggaggaccc gatcaacgcc acgggtgtgga agctccagcc cacagccggc ctccaggacc 840
tgcacatcca tccccggcag gaggaggagc agagcggatg catggagtac tcgggtgtgc 900
tgcctcgaac actcttccag aggacgaaag gcccggcgg ggaggcttag aagagactcc 960
tcctgggtgga cttagcgcagc caagccctgt tccaggacaa gaattccagg caagtctgg 1020
gtgagaaggt ctggggatt gtggtagaca acaccaaagt agccaaccc acggagcccg 1080
tggtagctcac ctcccagcac cagctacagc cgaagaatgt gactctgcaatgtgtttct 1140
gggttgaaga ccccacattt agcagccccc ggcattggag cagtgtggg tgtgagaccg 1200
tcaggagaga aacccaaaca tcctgttctt gcaaccactt gacctacttt gcagtgtga 1260
tggtagcttc ggtggaggtg gacggcgtgc acaagcaacta cctgagccct ctctccatcg 1320
tgggtgtgt ctgtctgtcc ctggcctggc ttgtcaccat tgccgcctac ctctgttcca 1380
gggtgccccctt gccgtgcagg agggaaaccc tcgttccatc catcaagggt cacatgaacc 1440
tgctgttgcg cgttcttctg ctggacacga gtttctgtc cagcgagccg gtggccctga 1500
caggtctga ggctggctgc cgagccagtg ccattttctt gcaacttctcc ctgctcacct 1560
gcctttctgt gatgggcctc gagggttaca accttaccc acgtgtggg gagggttttg 1620
gcacccatgt ccctggctac ctactcaagc tgagccat gggctggggc ttccccatct 1680
ttctgggtgac gctgggtggcc ctgggtggatg tggacaacta tggcccccattt atcttggctg 1740
tgcataaggac tccagaggcc gtcatctacc ctccatgtg ctggatccgg gactccctgg 1800
tcagctacat caccacccgt ggccttcca gcctgggtt tctgttcaac atggccatgc 1860
tagccacccat ggtgggtgcag atcctggccg tgegcacccca caccaaaaag tggtcacatg 1920
tgctgacact gctggccctc agcctggtcc ttggcctgccc ctggcccttg atcttcttct 1980
cctttgttcc tggcacccctc cagcttgcg tcctcttactt tttcagcattt atcaccctcc 2040
tccaaggctt cctcatcttc atctggtaact ggtccatgcg gtcgcaggcc cgggggtggcc 2100
cctcccccctt gaagagcaac tcagacagcg ccaggctccc catcagctcg ggcagcacct 2160
cggtccagccg catctaggcc tccagccac ctgcccattt gatgaaggcag agatgcggcc 2220
tcgttgcacca ctggcctgtgg ccccccggcc cggcccgagcc ccaggccagtg cagccgcaga 2280
ctttggaaag cccaaacgacc atggagagat gggccgttgc catgggtggac ggactcccg 2340
gctgggcttt tgaattggcc ttggggacta ctggccttc actcagctcc cacgggactc 2400
agaagtgcgc cgccatgtg ccttagggtac tgtccccaca tctgtcccaa cccagctgga 2460

ggcctggctct ctccttataa cccctggcc cagccctcat tgctgggggc caggccttgg 2520
 atcttggggg tctggcacat ccttaatctt gtgcctctgc ctggacaga aatgtggctc 2580
 cagttgcctc gtctctctgt gtcaccctgt gggcacatctg cttctctgtt cattttttaacc 2640
 tcaggtggca cccaggggca atggggccca gggcacatctt tcagggccag agccctggcg 2700
 gaggagaggg ccttggccag gacacagca gcagctcgcc tacctctgag cccaggcccc 2760
 ctccctccct cagcccccca gtcctccctc catcttccctt ggggttctcc ttcttccca 2820
 ggccttcctt gtccttcgt tcacagctgg gggtccccga ttccaatgct gtttttttggg 2880
 gagttgttcc caggagctgc ctgggtctg ctgttaatgt ttgtctactg cacaaggctc 2940
 ggcctgcctt tgagccagc tcggtaaccga tgcgtggctt gggctaggtc cctctgtcca 3000
 tctgggcctt tgtatgagct gcattggcctt tgctcaccctt gaccaagcac acgcctcaga 3060
 gggccctca gccttcctg aagccctttt gtggcaagaa ctgtggacca tgccagtccc 3120
 gtctgggttc catccccca ctccaaggac tgagactgac ctccctctgg gacactggcc 3180
 tagagcctga cacttcctta agaggttctc tccaagcccc caaatagctc caggccctt 3240
 cggccgcccc tcatggtaa ttctgtccaa caaacacaca cgggttagatt gctggctgt 3300
 ttaggggtt agggacacag atgaccgacc tggtcactcc tcctgccaac attcagtctg 3360
 gtatgtgagg cgtgcgtgaa gcaagaactc ctggagctac agggacaggg agccatcatt 3420
 cctgcctggg aatcctggaa gacttcctgc aggagtcage gttcaatctt gaccttgaag 3480
 atgggaagga tggttttttt acgttaccaat tcttttgc tttgatatta aaaagaagta 3540
 catgttcat ttagagaatt tggaaactgt agaagagaat caagaagaaa aataaaaaatc 3600
 agctgttga atcaccttagc aaaaaaaaaaa aa 3632

<210> 10

<211> 714

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2763296CB1

<400> 10

gggagccctcc cacgctctcc agctcactcg gcaggcagcg gggaccaggc ctggcagggtt 60
 aagcctctgg gggtggtatcc tggaaagggtt tccagccgccc tggccctcgat tgggaccctc 120
 caccctggcag cagggtgtga ctcccaagat tgactccgtc ggaggaaaat gactccccag 180
 tcgctgtgc agacacact gttctgtct agtctgtct tcctggtcca aggtgcccac 240
 ggcaggggcc acagggaaaga cttegtctc tgcaggccagc ggaaccagac acacaggagc 300
 agcctccact actactggtc catgcggctg caggccccggg gtggcccttc ccctctgaag 360
 agcaactcag acagcgccag gtcctccatc agtcggca gcacctcgat cagccgcata 420
 taggccttca gcccacctgc ccatgtatgt aagcagagat gcccgtcgat cgacactgc 480
 ctgttagcccc cgaacccggc ccagccccag gccagtaagc cgccagacttt agaaaggcca 540
 acgaccatgg agagatgggc ctggccatg gtggacggac tccggggctg ggcttttag 600
 attggcttag gggctactcg gtcctcactc agtcccacg ggactcaaga atgcggcgcc 660
 atgctgcctt aggtactgtc cccacatctg tcccaaccca gctggaggcc tgg 714

<210> 11

<211> 1035

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> 1029

<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte ID No: 3367641CB1

<400> 11
agactggctc tgagctcctg agctcagggtg atccacctgc cttagcctcc caaaaatgctg 60
ggattacagg cgtgagccac tgcacccggc cgccatgggg tttaaatttc tacaagaga 120
ctggggatgg ggatgtggag cttcctgtcc atccctcacc tcagcactgt tcactgtgtc 180
caaatctaag atgccagcca gagaactaag atcaagtgtg gtcttcagga atggcttgct 240
ccagctcgag catagcctgt gcagtgcac ttaagatTT ttccaaaaga tactaagcag 300
caggccccat gttggggagt caggatgggg gtctctgccc tgacagaact cacatcttg 360
ggaggtgact ccctgtcctg ttgtgttcca gacatcaacg agtgtgcac accgtcgaaa 420
gtgtcatgcg gaaaattctc ggactgctgg aacacagagg ggagctacga ctgcgtgtc 480
agcccccggat atgagctgt ttctggggca aaaacattca agaatgagag cgagaacacc 540
tgtcaaggta agaaccaccc cacgtcctt gacttccat ccatgaggtt tggggtcacc 600
agagccattc tggcagcattc cagagagcag ggccttgggtt gttagggtcag tgcctggcgt 660
ctgagatggg acaggtgcac atgtacctac cccaccaccc cagagaggca gagtgtatgt 720
aggattaagg gctgactctg ggaccagctg cctggccaa atcctgataaa ccaagtgtat 780
gacgttggcc aggatactta ttccagaaacc gctgtccaca tgtgagatgt gataacgtta 840
tcatgtttgt aattccagtg ctttggggagg ctgaggcggg aggatctt gaggccagga 900
gttcaagacc agcccatagc aacatagcaa gacccaccc ctataaaagg ggaggcgaag 960
gcaggttagat tgttttagcc caggggttcg agaccagcat gagcaagcgg gtgaaaccca 1020
tattatnng ttgca 1035

<210> 12
<211> 1002
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 866390CB1

<400> 12
atctggtaga actcagctgt gaattcctgt gatcctggc tttttctggc cagggaaatg 60
gaaagcgaga acagaacagt gataagagaa ttcatcctcc ttggctctgac ccagtctcaa 120
gatattcagc tcctggctt tttgtcttagtt ttaatattct acttcatcat cttcccttgg 180
aattttctca ttattttcac cataaaagtca gaccctggc tcacagcccc cctctatttc 240
tttctggca acttggcctt cctggatgca tcctacttct tcattgtggc tccccggatg 300
ttgggtgact tcctctctgc gaagaagata atctcctaca gaggctgcat cactcagctc 360
tttttcttgc acttccttgg aggagggggag ggattactcc ttgttgtgat ggccctttgac 420
cgctacatcg ccacatcgccg gctctgcac tatctactg tcatgaaccc tagaacctgc 480
tatgeaatga tttttggctt gttggcttggg ggttttgcactccattat ccaggtggtc 540
ctcatcctcc gcttgcctt ttgtggccca aaccagctgg acaacttctt ctgtgtatgtc 600
ccacaggtca tcaagctggc ctgcacccgac acatttggg tggagctt gatggcttc 660
aacagtggcc tttttttttt ctgtgtctt ctggggcttc tggccttcca tgcagtcatt 720
ctttgtcgca tacgagggtc ttcttctgag gcaaaaaaca aggccatgtc cacgtgcattc 780
accatataca ttgttatatt ctcatgttt ggacctggca tcttcatcta cacgcgcccc 840
ttcagggttt tccaggtga caaggtggtt tctctcttcc acacagtgtat ttttcctttg 900
ttgaatcctg tcatttatac ctttcgcac caggaagtga aagttccat gaaaaagggtg 960
ttaataagc acatagcctg aaaaaaaa aa 1002

<210> 13
<211> 1302
<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 998550CB1

<400> 13

tggacaaaagg gggcacacaca ttccttccat acgggtgagc ctctacactgc ctgggtgtgg 60
tcacagttca gcttcttcat gatggtgat cccaatggca atgaatccag tgctacatac 120
ttcatctaa taggcctccc tggtttagaa gaggctcagt tctgggttggc cttccattg 180
tgctccctct accttattgc tggcttaggt aacttgacaa tcatctacat tggcgact 240
gagcacagcc tgcacatggcc catgtatata tttctttgc tgccttcagg cattgacatac 300
ctcatctcca cctcatccat gcccataatg ctggccatct tctgggttcaa ttccactacc 360
atccagtttgc atgcttgc tgcacatggc tttggccatcc actccttatac tggcatggaa 420
tccacagtgc tgctggccat ggcttttgc cgctatgtgg ccatctgtca cccactgcgc 480
catgcccacag tacttacgtt gctctgtgc accaaaattt gttgggttgc tgggtgcgg 540
ggggctgcac tgcacatggcc cttctgttc ttcacatcaagc agctggccctt ctggccgtcc 600
aatatcttt cccatctca ctgcctacac caagatgtca tgaagctggc ctgtatgtat 660
atccgggtca atgtcgctca tggccttatac gtcacatcatct cccatctgtgg cctggactca 720
cttctcatct ccttctcata tctgcttatt cttaaagactg tgggtggctt gacacgtgaa 780
gcccagggca aggcatattgg cacttgcgtc tctcatgtgt gtgctgtgtt catattctat 840
gtaccttca ttggattgtc catgggtgc tgccttagca agcggcgtga ctctccgtg 900
cccgcatct tggccaaat tcatctgtg gttccctctg tgctcaaccc aattgtctat 960
ggagtgaaga caaaggagat tgcacagcgc atccttcgac tttccatgt gcccacacac 1020
gttcacagac ccttaggtgc agtgcataaa cttctttcc attcagagtc ctctgattca 1080
gattttaaatg ttaacatttt ggaagacatg attcagaaaaaaaatttctt taataaaaaat 1140
acaactcaga tccttcaaat atgaaactgg ttgggaatc tccatttttt caatattatt 1200
ttcttcttttgc tttcttgc acatataatt attaataaccc tgcacttagttt gttgtttag 1260
gttattact ttccattttca ccatgcgtc caaatctaaa ct 1302

<210> 14

<211> 1136

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1547873CB1

<400> 14

aagagagac tcatatttcag aggacaggaa gacaattcca gaccattggc ccacagaggt 60
gtcacctgtc ttctgtgtc tgggtcagcc tcaactgcggcc ctgacttacc cctgccccca 120
ggaaaaatgt cacagctgac atttggagca tagcagtcga tgcacatggcc acagacactg 180
cctgagactc acggccctgg agaaacgcgc attccttat tttccatgtgg tgaccaggag 240
gcacagaacc caaacctggt atctcagetc tgcgtgttct ttcttcaaaa tgacacaaat 300
gaaaccatac atatgcagat gagcatggca gtggacacgc agggccctgccc cttgaatatac 360
attggcccccaggctgtgtc ggtctccctc tgggtgggttct tattgaatgg cactgtcttc 420
tggctgtttt gctgtggggc cacgaatccc tacatggtat acatcttcca cctgggtcgct 480
gtgcacgtga tctatcttttgc tgcgtggca gtgggttct tacagattta ctctgtggct 540
gaagacatgg tcagttttgtt ccaccatggt gttttaaaata atgtatagtc tcccttttaag 600
gtgtatatttgc tcaaaaact ttttttaatg atttcttat tgcgttatttca aattatgtgt 660
acatgttact caaaaacaaag aagatttcatt caagtctgtat attaaaatatac atgggttaccc 720
cttaattccc atactgcatac atttgcgttgc tcatgtctca tgcacatggcc atttcaggag 780
catgtttaagg aggtgtcaga tgcacatggcc gggcatgaga cttctgtgg acaggtctt 840
gtctctccac ttctgttagc tctccatctca taagcagatt agtccctttt cagttgtcg 900

agctgcagat acaatttcca acatttcat tggtaactta cttcggttat gatTTaaaa 960
 acatgttatt gtcaagcaca ccgttttat aaatttgaat ttgtcagtct tttatgctt 1020
 atgtatttagt ttatgattaa agaagtctta cctacatgag ataatataaa atgaatctat 1080
 gctgtcttat actattaaat tatttcattt gtaaaatgt aaaagaaaaaaa aaaaaaa 1136

<210> 15
 <211> 1185
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2048484CB1

<400> 15
 tgcatcaagt cccaggcaag cccaggagtt gacatttctc tgcccagcca tgggcctcac 60
 cctgctcttgc tgcgtctcc tgggactaga aggtcaggc atagttggca gcctccctga 120
 ggtgctgcag gcacccgtgg gaagctccat tctggcag tgccactaca ggctccagga 180
 tgtcaagct cagaagggtgt ggtgcgggtt cttgcggag gggtgccagc ccctgggtgc 240
 ctcagctgtg gatcgcagag ctccagcggg caggcgtacg tttctcacag acctgggtgg 300
 gggcctgctg cagggtggaaa tggttaccct gcaggaaagag gatgctggcg agtatggctg 360
 catggtggat gggcccgagg gggcccagat tttgcacaga gtctctctga acatactgcc 420
 cccagagggaa gaagaagaga cccataagat tggcagtctg gctgagaacg cattctcaga 480
 ccctgcaggc agtgcacacc ctttggaaacc cagccaggat gagaagagca tccccttgat 540
 ctgggggtgct gtgtctctgg taggtctgtct ggtggcagcg gtgggtctgt ttgctgtat 600
 ggccaagagg aaacaagaat ccctcctcag tggccacca cgtcagtgc tctggaccgg 660
 ctgctgaatt gccttggat gtaccacaca ttaggcttga ctcaccaccc tcatttgaca 720
 ataccaccta caccagccta cctcttgatt ccccatcagg aaaaccttca ctcccagctc 780
 catcctcatt gccccctcta cctcctaagg tcctggtctg ctccaagcct gtgacatatg 840
 ccacagtaat ctccccggga gggacaagg gtggagggac ctcgtgtggg ccagccca 900
 atccaccta caatcagact ccatccagct aagctgtca tcacacttta aactcatgag 960
 gaccatccct aggggtctg tgcattccatc cagccagctc atgccctagg atccttagga 1020
 tatctgagca accaggact ttaagatcta atccaatgtc ctaactttac tagggaaagt 1080
 gacgctcaga catgactgag atgtcttggg gaagacctcc ctgcacccaa ctccccact 1140
 gtttcttcta ccattacaca ctggctaaa taaaccctaa taatg 1185

<210> 16
 <211> 2171
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5554170CB1

<400> 16
 cccacgcgtc cgccccacgcg tccgcgggt cctgcgcgtc ccggactgag gtggcgcccc 60
 tggggccggac ggcgggtgtcc cggcggtggcg ggaagccggc actggagcgg gagegcactg 120
 ggccgcgggac cgggaggcgc agggaccgga cggctcccgta gtgcggccacc tgacggtacc 180
 gagagggcgg cgcggccacc agcagagccg tccggccac tccctggga tctgacttgg 240
 ctcttgcgggt cggggcacc gtgaaggccct ggggtgtgcg tggctctcc tggtaggcgc 300
 ctttccccgg cgtccggctt ggggtgggtgg tggcggtgac tccagccccg cctctccctg 360
 gagaggaggc ctccactcgc tccctcgcc tccctccctg gggccgcgc gactcggcc 420
 ggcttccctgc tccctgcct gccggccggc cccgtggctt gaagaagtct tcacttccca 480
 ggagagccaa agcgtgtctg gcccctagggtg ggaaaagaac tggctgtgac ctttgcctg 540

accttggagg gcccagccctt gggctgaatg gcagcaccca cgcccgccccg tccggtgctg 600
 acccacctgc tggtggtctct ctccggcatg ggctctggg ctgcggtaa tgggatctgg 660
 gtggagctac ctgtgggtggt caaagagctt ccagagggtt ggagccctccc ctcttaacgtc 720
 tctgtgtttt tggctctggg gaacctgggt ctgctgggt tgaccctctg gaggaggctg 780
 gcccagggaa aggacgagca ggtccccatc cgggtggtgc aggtgtctggg catgggtggc 840
 acagccctgc tggcctctct gtggcaccat gtggcccccag tggcaggaca gttgcattct 900
 gtggcctct tagcaactggc ctttgtctg gcactggcat gctgtgcctc gaatgtca 960
 ttccctccct tcttggcca cctggccacct cgcttcttac ggtcattctt cctgggtcaa 1020
 ggcctgagtg ccctgtgcc ctgcgtctg gccctagtgc agggtgtggg ccgcctcgag 1080
 tgcccgcag cccccatcaa cggcaccctt ggccccccgc tcgacttctt tgagcgttt 1140
 cccgcacca ccttcttgc ggcactgact gccctctgg tcgcttcagc tgctgccttc 1200
 cagggttttc tgctgtgtt gcccacca ccatctgtac ccacaggggg gtaggatca 1260
 ggcctccagg tgggagcccc aggacgagag gaagagggtgg aagagtcctc accactgcaa 1320
 gagccaccaa gccaggcagc aggacccacc cctggccagg accctaaggc ctatcagett 1380
 ctatcagccc gcagtgcctg cctgctggc ctgtggccg ccaccaacgc gctgaccaat 1440
 ggcgtgtgc ctgcgtgca gagttttcc tgcttaccct acgggcgtct ggcctaccac 1500
 ctggctgtgg tgctggcag tgctgccaat cccctggct gcttcttggc catgggtgtg 1560
 ctgtgcaggt ccttggcagg gctggggcggc ctctctctgc tggcgtgtt ctgtgggggc 1620
 tacctgtatgg cgctggcagt cctgagcccc tgccccc tggcgtggc ac ctcggcgggg 1680
 gtggtcctcg tggcgtgtc gtgggtgtctg tgcttggcg tgcttcttcta cgtgaagggtg 1740
 gcagccagct ccctgtgca tggcggggggc cggccggcat tgctggcago cggcgtggcc 1800
 atccagggtgg gctctctgtct cggcgctgtt gctatgttcc ccccgaccag catctatcac 1860
 gtgttccaca gcagaaagga ctgtgcagac ccctgtgact cctgagccctg ggcagggtggg 1920
 gaccccgctc cccaacaccc ctgtttccct caatgtgcc accatgcctg agtgcctgca 1980
 gcccaggagg cccgcacacc ggtacactcg tggacaccta cacactccat aggagatct 2040
 ggcttccag ggtggcaag ggcaaggagc aggctggag ccaggacca gtggggctg 2100
 tagggtaagc ccctgagcct gggacctaca tgtgggttgc gtaataaaaac atttgttattt 2160
 aaaaaaaaaa a 2171

<210> 17

<211> 1014

<212> PRT

<213> Homo sapiens

<300>

<308> GenBank ID No: g2117161

<400> 17

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Val | Phe | Ser | Val | Arg | Gln | Cys | Gly | His | Val | Gly | Arg | Thr | Glu |
| 1 | | | | | | | | | | | | | | 15 |
| Glu | Val | Leu | Leu | Thr | Phe | Lys | Ile | Phe | Leu | Val | Ile | Ile | Cys | Leu |
| | | | | | | | | | | | | | | 30 |
| His | Val | Val | Leu | Val | Thr | Ser | Leu | Glu | Glu | Asp | Thr | Asp | Asn | Ser |
| | | | | | | | | | | | | | | 45 |
| Ser | Leu | Ser | Pro | Pro | Ala | Lys | Leu | Ser | Val | Val | Ser | Phe | Ala | |
| | | | | | | | | | | | | | | 60 |
| Pro | Ser | Ser | Asn | Glu | Val | Glu | Thr | Thr | Ser | Leu | Asn | Asp | Val | Thr |
| | | | | | | | | | | | | | | 75 |
| Leu | Ser | Leu | Leu | Pro | Ser | Asn | Glu | Thr | Glu | Lys | Thr | Lys | Ile | Thr |
| | | | | | | | | | | | | | | 90 |
| Ile | Val | Lys | Thr | Phe | Asn | Ala | Ser | Gly | Val | Lys | Pro | Gln | Arg | Asn |
| | | | | | | | | | | | | | | 105 |
| Ile | Cys | Asn | Leu | Ser | Ser | Ile | Cys | Asn | Asp | Ser | Ala | Phe | Phe | Arg |
| | | | | | | | | | | | | | | 120 |
| | | | | | | | | | | | | | | |
| 110 | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |
| 115 | | | | | | | | | | | | | | |

Gly Glu Ile Met Phe Gln Tyr Asp Lys Glu Ser Thr Val Pro Gln
 125 130 135
 Asn Gln His Ile Thr Asn Gly Thr Leu Thr Gly Val Leu Ser Leu
 140 145 150
 Ser Glu Leu Lys Arg Ser Glu Leu Asn Lys Thr Leu Gln Thr Leu
 155 160 165
 Ser Glu Thr Tyr Phe Ile Met Cys Ala Thr Ala Glu Ala Gln Ser
 170 175 180
 Thr Leu Asn Cys Thr Phe Thr Ile Lys Leu Asn Asn Thr Met Asn
 185 190 195
 Ala Cys Ala Ala Ile Ala Ala Leu Glu Arg Val Lys Ile Arg Pro
 200 205 210
 Met Glu His Cys Cys Cys Ser Val Arg Ile Pro Cys Pro Ser Ser
 215 220 225
 Pro Glu Glu Leu Gly Lys Leu Gln Cys Asp Leu Gln Asp Pro Ile
 230 235 240
 Val Cys Leu Ala Asp His Pro Arg Gly Pro Pro Phe Ser Ser Ser
 245 250 255
 Gln Ser Ile Pro Val Val Pro Arg Ala Thr Val Leu Ser Gln Val
 260 265 270
 Pro Lys Ala Thr Ser Phe Ala Glu Pro Pro Asp Tyr Ser Pro Val
 275 280 285
 Thr His Asn Val Pro Ser Pro Ile Gly Glu Ile Gln Pro Leu Ser
 290 295 300
 Pro Gln Pro Ser Ala Pro Ile Ala Ser Ser Pro Ala Ile Asp Met
 305 310 315
 Pro Pro Gln Ser Glu Thr Ile Ser Ser Pro Met Pro Gln Thr His
 320 325 330
 Val Ser Gly Thr Pro Pro Pro Val Lys Ala Ser Phe Ser Ser Pro
 335 340 345
 Thr Val Ser Ala Pro Ala Asn Val Asn Thr Thr Ser Ala Pro Pro
 350 355 360
 Val Gln Thr Asp Ile Val Asn Thr Ser Ser Ile Ser Asp Leu Glu
 365 370 375
 Asn Gln Val Leu Gln Met Glu Lys Ala Leu Ser Leu Gly Ser Leu
 380 385 390
 Glu Pro Asn Leu Ala Gly Glu Met Ile Asn Gln Val Ser Arg Leu
 395 400 405
 Leu His Ser Pro Pro Asp Met Leu Ala Pro Leu Ala Gln Arg Leu
 410 415 420
 Leu Lys Val Val Asp Asp Ile Gly Leu Gln Leu Asn Phe Ser Asn
 425 430 435
 Thr Thr Ile Ser Leu Thr Ser Pro Ser Leu Ala Leu Ala Val Ile
 440 445 450
 Arg Val Asn Ala Ser Ser Phe Asn Thr Thr Thr Phe Val Ala Gln
 455 460 465
 Asp Pro Ala Asn Leu Gln Val Ser Leu Glu Thr Gln Ala Pro Glu
 470 475 480
 Asn Ser Ile Gly Thr Ile Thr Leu Pro Ser Ser Leu Met Asn Asn
 485 490 495
 Leu Pro Ala His Asp Met Glu Leu Ala Ser Arg Val Gln Phe Asn
 500 505 510
 Phe Phe Glu Thr Pro Ala Leu Phe Gln Asp Pro Ser Leu Glu Asn
 515 520 525
 Leu Ser Leu Ile Ser Tyr Val Ile Ser Ser Ser Val Ala Asn Leu

| | | |
|-------------------------------------|-----------------------------|-----|
| 530 | 535 | 540 |
| Thr Val Arg Asn Leu Thr Arg Asn Val | Thr Val Thr Leu Lys His | |
| 545 | 550 | 555 |
| Ile Asn Pro Ser Gln Asp Glu Leu Thr | Val Arg Cys Val Phe Trp | |
| 560 | 565 | 570 |
| Asp Leu Gly Arg Asn Gly Gly Arg Gly | Gly Trp Ser Asp Asn Gly | |
| 575 | 580 | 585 |
| Cys Ser Val Lys Asp Arg Arg Leu Asn | Glu Thr Ile Cys Thr Cys | |
| 590 | 595 | 600 |
| Ser His Leu Thr Ser Phe Gly Val Leu | Leu Asp Leu Ser Arg Thr | |
| 605 | 610 | 615 |
| Ser Val Leu Pro Ala Gln Met Met Ala | Leu Thr Phe Ile Thr Tyr | |
| 620 | 625 | 630 |
| Ile Gly Cys Gly Leu Ser Ser Ile Phe | Leu Ser Val Thr Leu Val | |
| 635 | 640 | 645 |
| Thr Tyr Ile Ala Phe Glu Lys Ile Arg | Arg Asp Tyr Pro Ser Lys | |
| 650 | 655 | 660 |
| Ile Leu Ile Gln Leu Cys Ala Ala Leu | Leu Leu Leu Asn Leu Val | |
| 665 | 670 | 675 |
| Phe Leu Leu Asp Ser Trp Ile Ala Leu | Tyr Lys Met Gln Gly Leu | |
| 680 | 685 | 690 |
| Cys Ile Ser Val Ala Val Phe Leu His | Tyr Phe Leu Leu Val Ser | |
| 695 | 700 | 705 |
| Phe Thr Trp Met Gly Leu Glu Ala Phe | His Met Tyr Leu Ala Leu | |
| 710 | 715 | 720 |
| Val Lys Val Phe Asn Thr Tyr Ile Arg | Lys Tyr Ile Leu Lys Phe | |
| 725 | 730 | 735 |
| Cys Ile Val Gly Trp Gly Val Pro Ala | Val Val Val Thr Ile Ile | |
| 740 | 745 | 750 |
| Leu Thr Ile Ser Pro Asp Asn Tyr Gly | Leu Gly Ser Tyr Gly Lys | |
| 755 | 760 | 765 |
| Phe Pro Asn Gly Ser Pro Asp Asp Phe | Cys Trp Ile Asn Asn Asn | |
| 770 | 775 | 780 |
| Ala Val Phe Tyr Ile Thr Val Val Gly | Tyr Phe Cys Val Ile Phe | |
| 785 | 790 | 795 |
| Leu Leu Asn Val Ser Met Phe Ile Val | Val Val Leu Val Gln Leu Cys | |
| 800 | 805 | 810 |
| Arg Ile Lys Lys Lys Gln Leu Gly Ala | Gln Arg Lys Thr Ser | |
| 815 | 820 | 825 |
| Ile Gln Asp Leu Arg Ser Ile Ala Gly | Leu Thr Phe Leu Leu Gly | |
| 830 | 835 | 840 |
| Ile Thr Trp Gly Phe Ala Phe Ala | Trp Gly Pro Val Asn Val | |
| 845 | 850 | 855 |
| Thr Phe Met Tyr Leu Phe Ala Ile Phe | Asn Thr Leu Gln Gly Phe | |
| 860 | 865 | 870 |
| Phe Ile Phe Ile Phe Tyr Cys Val Ala | Lys Glu Asn Val Arg Lys | |
| 875 | 880 | 885 |
| Gln Trp Arg Arg Tyr Leu Cys Cys Gly | Lys Leu Arg Leu Ala Glu | |
| 890 | 895 | 900 |
| Asn Ser Asp Trp Ser Lys Thr Ala Thr | Asn Gly Leu Lys Lys Gln | |
| 905 | 910 | 915 |
| Thr Val Asn Gln Gly Val Ser Ser Ser | Asn Ser Leu Gln Ser | |
| 920 | 925 | 930 |
| Ser Ser Asn Ser Thr Asn Ser Thr Thr | Leu Leu Val Asn Asn Asp | |
| 935 | 940 | 945 |

| | | |
|---|------|------|
| Cys Ser Val His Ala Ser Gly Asn Gly Asn Ala Ser Thr Glu Arg | | |
| 950 | 955 | 960 |
| Asn Gly Val Ser Phe Ser Val Gln Asn Gly Asp Val Cys Leu His | | |
| 965 | 970 | 975 |
| Asp Phe Thr Gly Lys Gln His Met Phe Asn Glu Lys Glu Asp Ser | | |
| 980 | 985 | 990 |
| Cys Asn Gly Lys Gly Arg Met Ala Leu Arg Arg Thr Ser Lys Arg | | |
| 995 | 1000 | 1005 |
| Gly Ser Leu His Phe Ile Glu Gln Met | | |
| 1010 | | |

<210> 18
 <211> 287
 <212> PRT
 <213> Homo sapiens

<300>
 <308> GenBank ID No: g3342740

| | | | |
|---|-----|-----|----|
| <400> 18 | | | |
| Met Gly Gly Arg Val Phe Leu Ala Phe Cys Val Trp Leu Thr Leu | | | |
| 1 | 5 | 10 | 15 |
| Pro Gly Ala Glu Thr Gln Asp Ser Arg Gly Cys Ala Arg Trp Cys | | | |
| 20 | 25 | 30 | |
| Pro Gln Asp Ser Ser Cys Val Asn Ala Thr Ala Cys Arg Cys Asn | | | |
| 35 | 40 | 45 | |
| Pro Gly Phe Ser Ser Phe Ser Glu Ile Ile Thr Thr Pro Asp Ile | | | |
| 50 | 55 | 60 | |
| Asn Glu Cys Ala Thr Leu Ser Lys Val Ser Cys Gly Lys Phe Ser | | | |
| 65 | 70 | 75 | |
| Asp Cys Trp Asn Thr Glu Gly Ser Tyr Asp Cys Val Cys Ser Pro | | | |
| 80 | 85 | 90 | |
| Gly Tyr Glu Pro Val Ser Gly Ala Lys Thr Phe Lys Asn Glu Ser | | | |
| 95 | 100 | 105 | |
| Glu Asn Thr Cys Gln Asp Val Asp Glu Cys Gln Gln Asn Pro Arg | | | |
| 110 | 115 | 120 | |
| Leu Cys Lys Ser Tyr Gly Thr Cys Val Asn Thr Leu Gly Ser Tyr | | | |
| 125 | 130 | 135 | |
| Thr Cys Gln Cys Leu Pro Gly Phe Lys Asp Val Asn Glu Cys Thr | | | |
| 140 | 145 | 150 | |
| Ser Gly Gln Asn Pro Cys His Ser Ser Thr His Cys Leu Asn Asn | | | |
| 155 | 160 | 165 | |
| Val Gly Ser Tyr Gln Cys Arg Cys Arg Pro Gly Trp Gln Pro Ile | | | |
| 170 | 175 | 180 | |
| Pro Gly Ser Pro Asn Gly Pro Asn Asn Thr Val Cys Glu Asp Val | | | |
| 185 | 190 | 195 | |
| Asp Glu Cys Ser Ser Gly Gln His Gln Cys Asp Ser Ser Thr Val | | | |
| 200 | 205 | 210 | |
| Cys Phe Asn Thr Val Gly Ser Tyr Ser Cys Arg Cys Arg Pro Gly | | | |
| 215 | 220 | 225 | |
| Trp Lys Pro Arg His Gly Ile Pro Asn Asn Gln Lys Asp Thr Val | | | |
| 230 | 235 | 240 | |
| Cys Glu Asp Met Thr Phe Ser Thr Trp Thr Pro Pro Pro Gly Val | | | |

| | | |
|---|-----|-----|
| 245 | 250 | 255 |
| His Ser Gln Thr Leu Ser Arg Phe Phe Asp Lys Val Gln Asp Leu | | |
| 260 | 265 | 270 |
| Gly Arg Asp His Leu Ser Ser Phe Ala Val Leu Met Ala His Tyr | | |
| 275 | 280 | 285 |
| Asp Val | | |



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

| | | | | | | | | | | | | | | | | | | | | | | | | |
|--|---------------------------|----|---|------------|---------------------------|----|------------|---------------------------|----|------------|------------------------|----|----|------------------|----------|---------------------------|----|------------------|----------|---------------------------|----|------------------|----------|------------------------|
| (51) International Patent Classification ⁷ : C12N 15/12, 5/10, 1/21, C07K 14/705, 14/72, 16/28, A61K 38/17, C12Q 1/68 | | A3 | (11) International Publication Number: WO 00/20590 | | | | | | | | | | | | | | | | | | | | | |
| | | | (43) International Publication Date: 13 April 2000 (13.04.00) | | | | | | | | | | | | | | | | | | | | | |
| <p>(21) International Application Number: PCT/US99/23317</p> <p>(22) International Filing Date: 6 October 1999 (06.10.99)</p> <p>(30) Priority Data:</p> <table> <tr> <td>09/167,219</td> <td>6 October 1998 (06.10.98)</td> <td>US</td> </tr> <tr> <td>60/172,211</td> <td>6 October 1998 (06.10.98)</td> <td>US</td> </tr> <tr> <td>60/133,585</td> <td>11 May 1999 (11.05.99)</td> <td>US</td> </tr> </table> <p>(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications</p> <table> <tr> <td>US</td> <td>09/167,219 (CIP)</td> </tr> <tr> <td>Filed on</td> <td>6 October 1998 (06.10.98)</td> </tr> <tr> <td>US</td> <td>60/172,211 (CIP)</td> </tr> <tr> <td>Filed on</td> <td>6 October 1998 (06.10.98)</td> </tr> <tr> <td>US</td> <td>60/133,585 (CIP)</td> </tr> <tr> <td>Filed on</td> <td>11 May 1999 (11.05.99)</td> </tr> </table> <p>(71) Applicant (for all designated States except US): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).</p> <p>(72) Inventors; and</p> <p>(75) Inventors/Applicants (for US only): TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). AU-YOUNG, Janice [US/US]; 1419 Kains Avenue, Berkeley, CA 94702 (US). REDDY, Roopa [IN/US]; 1233 W. McKinley Drive, Sunnyvale, CA 94086 (US). CORLEY, Neil, C. [US/US];</p> | | | | 09/167,219 | 6 October 1998 (06.10.98) | US | 60/172,211 | 6 October 1998 (06.10.98) | US | 60/133,585 | 11 May 1999 (11.05.99) | US | US | 09/167,219 (CIP) | Filed on | 6 October 1998 (06.10.98) | US | 60/172,211 (CIP) | Filed on | 6 October 1998 (06.10.98) | US | 60/133,585 (CIP) | Filed on | 11 May 1999 (11.05.99) |
| 09/167,219 | 6 October 1998 (06.10.98) | US | | | | | | | | | | | | | | | | | | | | | | |
| 60/172,211 | 6 October 1998 (06.10.98) | US | | | | | | | | | | | | | | | | | | | | | | |
| 60/133,585 | 11 May 1999 (11.05.99) | US | | | | | | | | | | | | | | | | | | | | | | |
| US | 09/167,219 (CIP) | | | | | | | | | | | | | | | | | | | | | | | |
| Filed on | 6 October 1998 (06.10.98) | | | | | | | | | | | | | | | | | | | | | | | |
| US | 60/172,211 (CIP) | | | | | | | | | | | | | | | | | | | | | | | |
| Filed on | 6 October 1998 (06.10.98) | | | | | | | | | | | | | | | | | | | | | | | |
| US | 60/133,585 (CIP) | | | | | | | | | | | | | | | | | | | | | | | |
| Filed on | 11 May 1999 (11.05.99) | | | | | | | | | | | | | | | | | | | | | | | |
| <p>(1240 Dale Avenue #30, Mountain View, CA 94040 (US). GUEGLER, Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). GORGONE, Gina, A. [US/US]; 1253 Pinecrest Drive, Boulder Creek, CA 95006 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US).</p> <p>(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).</p> <p>(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p>Published <i>With international search report.</i></p> <p>(88) Date of publication of the international search report: 20 July 2000 (20.07.00)</p> | | | | | | | | | | | | | | | | | | | | | | | | |
| <p>(54) Title: G-PROTEIN COUPLED RECEPTOR PROTEINS</p> <p>1 M-----TPQSLLQTTLFLLSL-----LFLV 2763296 1 M-----TPQSLLQTTLFLLSL-----LFLV 1595762 1 MVF SVRQC'GHVGRTEEVVLTFKIPFLVIVICLHVVLV GI 2117161</p> <p>21 QGAHGRGHREDFR-----FCSQRNQTHRSS 2763296 21 QGAHGRGHREDFR-----FCSQRNQTHRSS 1595762 36 T SLEEDTDNSSLSPPPAAKLSVVSFAFSSNEVEETTS GI 2117161</p> <p>46 LHY-----FCSQRNQTHRSS 2763296 46 LHYKPTFDLRISIENSEEALTVHAPFFAAHPASR-----FCSQRNQTHRSS 1595762 71 LIN-DVTLSSLPSNETEKTKITIVKTFNAGVKPQR GI 2117161</p> <p>49 -----FCSQRNQTHRSS 2763296 80 -----SFPPDPRGLYHF-----CLYWNRH 1595762 105 NICNLSSICNDSAFFRGEIMFQYDKESTVTPQNQH GI 2117161</p> <p>49 -----FCSQRNQTHRSS 2763296 98 -----AGRLHLLYGKRD-----FLL 1595762 140 TNGTLTGVLSSLSELKRSLENKTLQTLSETYFIMCA GI 2117161</p> <p>49 -----FCSQRNQTHRSS 2763296 113 -----SDKASSLLC 1595762 175 TAEAQSTLNCTFTIKLNNTMNACAAIAALERVKIR GI 2117161</p> | | | | | | | | | | | | | | | | | | | | | | | | |
| <p>(57) Abstract</p> <p>The invention provides human G-protein coupled receptor proteins (GCRP) and polynucleotides which identify and encode GCRP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of GCRP.</p> | | | | | | | | | | | | | | | | | | | | | | | | |

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | | | |
|----|--------------------------|----|---------------------------------------|----|---|----|--------------------------|
| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| AT | Austria | FR | France | LU | Luxembourg | SN | Senegal |
| AU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav Republic of Macedonia | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | | | TR | Turkey |
| BG | Bulgaria | HU | Hungary | ML | Mali | TT | Trinidad and Tobago |
| BJ | Benin | IE | Ireland | MN | Mongolia | UA | Ukraine |
| BR | Brazil | IL | Israel | MR | Mauritania | UG | Uganda |
| BY | Belarus | IS | Iceland | MW | Malawi | US | United States of America |
| CA | Canada | IT | Italy | MX | Mexico | UZ | Uzbekistan |
| CF | Central African Republic | JP | Japan | NE | Niger | VN | Viet Nam |
| CG | Congo | KE | Kenya | NL | Netherlands | YU | Yugoslavia |
| CH | Switzerland | KG | Kyrgyzstan | NO | Norway | ZW | Zimbabwe |
| CI | Côte d'Ivoire | KP | Democratic People's Republic of Korea | NZ | New Zealand | | |
| CM | Cameroon | | | PL | Poland | | |
| CN | China | KR | Republic of Korea | PT | Portugal | | |
| CU | Cuba | KZ | Kazakhstan | RO | Romania | | |
| CZ | Czech Republic | LC | Saint Lucia | RU | Russian Federation | | |
| DE | Germany | LI | Liechtenstein | SD | Sudan | | |
| DK | Denmark | LK | Sri Lanka | SE | Sweden | | |
| EE | Estonia | LR | Liberia | SG | Singapore | | |

INTERNATIONAL SEARCH REPORT

Intern: Application No
PCT/US 99/23317

| A. CLASSIFICATION OF SUBJECT MATTER | | | | | |
|-------------------------------------|-----------|----------|----------|------------|-----------|
| IPC 7 | C12N15/12 | C12N5/10 | C12N1/21 | C07K14/705 | C07K14/72 |

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K A61K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ^a | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------------------|--|-----------------------|
| X | <p>DATABASE EMEST30 [Online] E.M.B.L. Databases Accession Number: C17798, 4 October 1996 (1996-10-04) FUJIWARA T ET AL: "Human placenta cDNA 5'-end GEN-553A03" XP002130338 99.8% identity in 556 bp overlap with SeqIdNo.9 abstract</p> <p>---</p> | 3-6 |
| P,X | <p>WO 99 46281 A (BAKER KEVIN P ;CHEN JIAN (US); GENENTECH INC (US); GURNEY AUSTIN () 16 September 1999 (1999-09-16) Human PRO1083: 100.0% identity in 693 aa overlap with SeqIdNo.1 / 99.9% identity in 3574 bp overlap with SeqIdNo.9</p> <p>---</p> <p>-/-</p> | 1-16,19 |

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

^a Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *&* document member of the same patent family

Date of the actual completion of the international search

11 February 2000

Date of mailing of the international search report

16. 05. 2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel: (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Lonnoy, O

INTERNATIONAL SEARCH REPORT

Intern: _____ Application No

PCT/US 99/23317

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|------------|--|-----------------------|
| P,X | WO 99 15551 A (GARVAN INST MED RES ;HERZOG HERBERT (AU)) 1 April 1999 (1999-04-01) TSR32: 99.9% identity in 693 aa overlap with SeqIdNo.1 / 99.7% identity in 2699 bp overlap with SeqIdNo.9 --- | 1-16,19 |
| A | EP 0 805 204 A (INST HORMON UND FORTPFLANZUNGS) 5 November 1997 (1997-11-05) --- | |
| A | WO 96 18651 A (ELLIS CATHERINE ELIZABETH ;SMITHKLINE BEECHAM CORP (US); BERGSMA D) 20 June 1996 (1996-06-20) --- | |
| A | STADEL J M ET AL: "Orphan G protein-coupled receptors: a neglected opportunity for pioneer drug discovery" TRENDS IN PHARMACOLOGICAL SCIENCES, GB, ELSEVIER TRENDS JOURNAL, CAMBRIDGE, vol. 18, no. 11, November 1997 (1997-11), page 430-437 XP004096215 ISSN: 0165-6147 ----- | |

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 99/23317

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:
Although claim 19 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. Claims Nos.: 17, 18, 20 because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

see additional sheet, invention 1.

Remark on Protest

The additional search fees were accompanied by the applicant's protest.
 No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 17,18,20

Claims 17, 18 and 20 refer to an agonist, an antagonist and to the use of an antagonist of the polypeptide of claim 1 without giving a true technical characterization. Moreover, no such compound is defined in the application. In consequence, the scope of said claims is ambiguous and vague, and their subject-matter is not sufficiently disclosed and supported (Art. 5 and 6 PCT). No search can be carried out for such speculative claims the wording of which is, in fact, a mere recitation of the results to be achieved.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Invention 1: claims 1-20 all partially
A substantially purified polypeptide comprising an amino acid sequence of SeqIdNo.1, a fragment of said polypeptide, a variant polypeptide having at least 90% sequence identity to said polypeptide; an isolated and purified polynucleotide encoding said polypeptide, a variant polynucleotide having at least 70% sequence identity to said polynucleotide, a polynucleotide hybridising to said polynucleotide, a polynucleotide having a sequence complementary to said polynucleotide; a detection method using said complementary polynucleotide; a polynucleotide comprising a sequence of SeqIdNo.9 or a fragment thereof, a polynucleotide having at least 70% sequence identity to said polynucleotide of SeqIdNo.9, a polynucleotide complementary to said polynucleotide of SeqIdNo.9; expression vector, host and recombinant method related thereto; a pharmaceutical composition comprising said polypeptide of SeqIdNo.1 or fragment of said polypeptide of SeqIdNo.1; an antibody to said polypeptide of SeqIdNo.1 or fragment of said polypeptide of SeqIdNo.1; an agonist to said polypeptide of SeqIdNo.1; an antagonist to said polypeptide of SeqIdNo.1; therapeutic application thereof

2. Inventions 2-7: claims 1-20 all partially
Idem as for subject 1 but limited to each of the polypeptide sequences as in SeqIdNo.3-8 and the corresponding polynucleotide sequences as in SeqIdNo:11-16, wherein respectively invention 2 is limited to SeqIdNo:3 and SeqIdNo.11, invention 3 is limited to SeqIdNo:4 and SeqIdNo.12, ..., invention 7 is limited to SeqIdNo.8 and SeqIdNo.16

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 99/23317

| Patent document cited in search report | Publication date | Patent family member(s) | | Publication date |
|--|------------------|-------------------------|------------|------------------|
| WO 9946281 | A 16-09-1999 | AU | 3072199 A | 27-09-1999 |
| | | AU | 3075099 A | 11-10-1999 |
| | | WO | 9947677 A | 23-09-1999 |
| | | AU | 1532499 A | 15-06-1999 |
| | | WO | 9927098 A | 03-06-1999 |
| | | AU | 3757099 A | 08-11-1999 |
| | | WO | 9954467 A | 28-10-1999 |
| | | AU | 1070399 A | 10-05-1999 |
| | | WO | 9920756 A | 29-04-1999 |
| ----- | ----- | ----- | ----- | ----- |
| WO 9915551 | A 01-04-1999 | AU | 9330098 A | 12-04-1999 |
| | | EP | 0973800 A | 26-01-2000 |
| ----- | ----- | ----- | ----- | ----- |
| EP 0805204 | A 05-11-1997 | DE | 19617940 A | 30-10-1997 |
| ----- | ----- | ----- | ----- | ----- |
| WO 9618651 | A 20-06-1996 | AU | 4522796 A | 03-07-1996 |
| | | EP | 0871669 A | 21-10-1998 |
| | | JP | 10511936 T | 17-11-1998 |
| | | ZA | 9510686 A | 03-07-1996 |
| ----- | ----- | ----- | ----- | ----- |